

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 7, 1999, 21:20:10 ; Search time 465.18 Seconds

(without alignments)
11875.397 Million cell updates/sec

Title: US-09-104-063-3

Perfect score: 1737

Sequence: 1 GAATTCGACTGCTGCTGGCGG.....CCGCCACGACTGGAATTC 1737

Scoring table: OLIGO_NUC

Searched: 579419 seqs, 1590154680 residues

Database :

GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_om:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_pl1:*
8: gb_pl2:*
9: gb_pl1:*
10: gb_pl2:*
11: gb_pl3:*
12: gb_pl4:*
13: gb_pl5:*
14: gb_pl6:*
15: gb_pl7:*
16: gb_pl8:*
17: gb_pl9:*
18: gb_pl10:*
19: gb_pl11:*
20: gb_pl12:*
21: gb_pl13:*
22: gb_pl14:*
23: gb_pl15:*
24: gb_pl16:*
25: gb_pl17:*
26: gb_pl18:*
27: gb_pl19:*
28: gb_pl20:*
29: gb_pl21:*
30: gb_pl22:*
31: gb_pl23:*
32: gb_pl24:*
33: gb_pl25:*
34: gb_pl26:*
35: gb_pl27:*
36: gb_pl28:*
37: gb_pl29:*
38: gb_pl30:*
39: gb_pl31:*
40: gb_pl32:*
41: gb_pl33:*
42: gb_pl34:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match Length	ID	Description
1	1737	100.0	1737 5	AR015970 Sequence

2	1737	100.0	1737 5	I13753 Sequence 4
3	1737	100.0	1737 5	I24455 Sequence 3
4	1642	94.5	1645 10	HSNPRFLA
5	1601	92.2	1664 9	HUMHMR89
6	1601	92.2	1664 14	G28514 human mRNA
7	1564	90.0	5161 11	AF005058 Homo sapi
8	1536	88.4	3733 10	HSXCR4
9	1513	87.1	8747 9	HSR24869 Homo sapi
10	1513	87.1	8747 11	AF032572
11	1278	73.6	1637 9	HUMSTSR
12	1245	71.7	1670 9	HUMGPCR
13	1108	63.8	1225 9	HUMNYRCA
14	1059	61.0	1059 11	AF025375
15	985	56.7	1087 11	PTU89798
16	341	19.6	1059 10	MM093111
17	341	19.6	1087 11	AF001928
18	336	19.3	1078 9	AF001928
19	313	18.0	1126 11	AF019378
20	313	18.0	1059 11	AF031089
21	239	13.8	1059 11	AF051906
22	239	13.8	1059 11	MM073740
23	131	7.5	1089 11	AB015943
24	77	4.4	1689 3	FCA9816
25	77	4.4	1170 3	FCU63558
26	77	4.4	1140 3	FCU92795
27	70	4.0	95 9	HUMNYR82
28	62	3.6	1280 3	BOVNYR
29	56	3.2	1223 12	MM059760
30	56	3.2	3366 12	MM065580
31	56	3.2	1362 12	RNU54791
32	56	3.2	1050 12	RNU90610
33	53	3.1	3581 5	E07381
34	53	3.1	995 42	AF002985
35	50	2.9	578 3	OMI38942
36	50	2.9	60 5	I88893
37	50	2.9	1877 12	D87747
38	50	2.9	495 36	S81785
39	49	2.8	3362 4	AF139986
40	49	2.8	2035 4	AF153205
41	49	2.8	1772 7	SACATATAS
42	49	2.8	711 8	AF069331
43	49	2.8	1207 8	AF094831
44	49	2.8	1845 10	HSR800467
45	49	2.8	1403 12	MMU132433
46	49	2.8	754 12	MUSCALTRA
47	49	2.8	987 36	AFCAATP
48	49	2.8	982 36	BMOTELRPT
49	48	2.8	1032 3	AF013214
50	48	2.8	1338 4	ECCD44
51	48	2.8	2374 4	XLA5870
52	48	2.8	1490 5	A41579
53	48	2.8	1474 5	A65341
54	48	2.8	1474 5	E06743
55	48	2.8	2638 7	ATRA0058
56	48	2.8	1425 7	ATGOX2353
57	48	2.8	1083 7	ATH10467
58	48	2.8	1586 7	BNPML3
59	48	2.8	1493 7	HVCHS2
60	48	2.8	1572 8	AF014053
61	48	2.8	2047 8	AF035815
62	48	2.8	1960 8	AF150881
63	48	2.8	893 8	ATU75199
64	48	2.8	1589 8	PTX13769
65	48	2.8	5503 10	HSCH16FAA
66	48	2.8	1887 10	HSGLCNACT
67	48	2.8	3319 10	HSNRRNPI
68	48	2.8	1746 10	HSR800208
69	48	2.8	3462 10	HSR800550
70	48	2.8	1373 10	HSY14040
71	48	2.8	185741 11	AC006203
72	48	2.8	1448 12	AF067728
73	48	2.8	1095 12	CATRANS
74	48	2.8	159319 34	HS40E16

75	48	2.8	944	36	BF283268	283268 B. floridiae	148	47	2.7	989	7	CPCRMT	Y08233 C. pteroides
76	48	2.8	8970	36	BMAMYNLTRR	U07847 Bombyx mori	149	47	2.7	1566	7	CRO7333	AJ007333 Catharant
77	48	2.8	1248	36	HMRAS1	X70839 H. magnipapili	150	47	2.7	1686	7	MECGP5	X77462 M. esculenta
78	48	2.8	2312	36	TG093689	U93689 Toxoplasma	151	47	2.7	1975	7	NTPOLPHOX	Y12501 Nicotiana t
79	48	2.8	1159	37	AF080131	AF080131 Drosophila	152	47	2.7	1577	7	OSCHSLKCN	X1811 O. sativa mR
80	47	2.7	931	37	AF054289	AF054289 Macropus	153	47	2.7	1739	7	PCU42442	Y02442 P. deltoideus
81	47	2.7	2022	3	BTDSPO	X95310 B. taurus mR	154	47	2.7	1305	7	PDCLALDHA	Z19568 P. deltoideus
82	47	2.7	4719	3	BRRNMXOR	X83508 B. taurus mR	155	47	2.7	572	7	PPY16958	PPY16958 P. deltoideus
83	47	2.7	3934	3	RABMEMGLX	L04504 Oryctolagus	156	47	2.7	554	7	PPY16960	PPY16960 P. deltoideus
84	47	2.7	732	3	S65585	S65585 refinol-bin	157	47	2.7	1528	7	PSAJ2155	PSAJ2155 P. deltoideus
85	47	2.7	2307	4	AF043493	AF043493 Gallus ga	158	47	2.7	1022	7	PSPIAA45R	PSPIAA45R P. deltoideus
86	47	2.7	3188	4	AF124396	AF124396 Dario rer	159	47	2.7	2051	7	SOCPKNA	X01355 R. olivaceus P
87	47	2.7	4230	4	AF124435	AF124435 Dario rer	160	47	2.7	2647	7	SOPRRKINB	Z30332 S. olivaceus
88	47	2.7	1584	4	CCRHO	U07445 Cyprinus ca	161	47	2.7	2532	7	TAU48227	U48227 Trilicium ae
89	47	2.7	4387	4	DRE5690	AI005690 Dario rer	162	47	2.7	2055	7	THIGPMR	Z11528 T. harzianum
90	47	2.7	3348	4	DRZP47POU	X07905 D. rerio mRN	163	47	2.7	1061	8	AF002226	AF002226 Nicotiana
91	47	2.7	1208	4	GSPMADPR	X83397 Gallus sp.	164	47	2.7	759	8	AF008214	AF008214 Musa acum
92	47	2.7	1584	4	S74449	S74449 rhodopsin l	165	47	2.7	853	8	AF049923	AF049923 Petunia x
93	47	2.7	1845	4	XLGYCE	Z13966 X. laevis mR	166	47	2.7	868	8	AF049925	AF049925 Petunia x
94	47	2.7	4435	4	XLMPTLG	Z97214 Xenopus lae	167	47	2.7	625	8	AF064552	AF064552 Apium gra
95	47	2.7	2997	4	XLMPTLG	X55813 X. laevis MP	168	47	2.7	1320	8	AF064552	AF064552 Apium gra
96	47	2.7	2470	4	XLU62807	U62807 Xenopus lae	169	47	2.7	957	8	AF071890	AF071890 Pinus ar
97	47	2.7	1239	5	A08456	A08456 Recombinant	170	47	2.7	1425	8	AF076252	AF076252 Arabidops
98	47	2.7	102	5	A08911	A08911 H. sapiens (171	47	2.7	1043	8	AF079185	AF079185 Arabidops
99	47	2.7	107	5	A08912	A08912 H. sapiens (172	47	2.7	605	8	AF123482	AF123482 Myosinphae
100	47	2.7	132	5	A08913	A08913 H. sapiens (173	47	2.7	1570	8	ATP2UR2	X74766 Arabidops
101	47	2.7	843	5	A12552	A12552 Ag169 clone	174	47	2.7	1809	8	ATP2UR2	U79733 Beta procum
102	47	2.7	1165	5	A12558	A12558 P. falciparu	175	47	2.7	1450	8	ATU53864	ATU53864 Arabidops
103	47	2.7	1078	5	A26498	A26498 Batley PSI	176	47	2.7	1748	8	BP079733	BP079733 Arabidops
104	47	2.7	1243	5	A31057	A31057 L. perenne m	177	47	2.7	1360	8	CTAJ5082	CTAJ5082 Arabidops
105	47	2.7	1748	5	A44314	A44314 Sequence 1	178	47	2.7	616	8	SFUS8050	SFUS8050 Arabidops
106	47	2.7	1141	5	A48220	A48220 Sequence 19	179	47	2.7	200000	9	AB019438	AB019438 Homo sapi
107	47	2.7	767	5	A48221	A48221 Sequence 20	180	47	2.7	100000	9	AP000081	AP000081 Homo sapi
108	47	2.7	1141	5	A57313	A57313 Sequence 19	181	47	2.7	102456	9	AP000083	AP000083 Homo sapi
109	47	2.7	1604	5	AR009628	AR009628 Sequence 20	182	47	2.7	100375	9	HS193B12	HS193B12 Homo sapi
110	47	2.7	568	5	AR012379	AR012379 Sequence	183	47	2.7	89328	9	HS196C22	HS196C22 Homo sapi
111	47	2.7	1172	5	AR013797	AR013797 Sequence	184	47	2.7	36669	9	HS433G19	HS433G19 Homo sapi
112	47	2.7	458	5	AR022283	AR022283 Sequence	185	47	2.7	4413	9	HS4012755	HS4012755 Homo sapi
113	47	2.7	1794	5	E02349	E02349 cDNA sequen	186	47	2.7	6571	9	HGMP220	HGMP220 Homo sapien
114	47	2.7	2051	5	E03168	E03168 Promoter of	187	47	2.7	1976	9	HGMP220	HGMP220 Homo sapien
115	47	2.7	916	5	E12747	E12747 Rhodnius pr	188	47	2.7	3380	10	AF002672	AF002672 Homo sapi
116	47	2.7	1868	5	E12806	E12806 cDNA encodi	189	47	2.7	43761	10	HS57B2	HS57B2 Homo sapi
117	47	2.7	766	5	E108319	E108319 Sequence 1	190	47	2.7	2530	10	HS57B2	HS57B2 Homo sapi
118	47	2.7	2549	5	I09499	I09499 Sequence 20	191	47	2.7	1598	10	HS57B2	HS57B2 Homo sapi
119	47	2.7	1172	5	I13099	I13099 Sequence 9	192	47	2.7	1606	10	HS57B2	HS57B2 Homo sapi
120	47	2.7	399	5	I18358	I18358 Sequence 13	193	47	2.7	3174	10	HS57B2	HS57B2 Homo sapi
121	47	2.7	1172	5	I25472	I25472 Sequence 3	194	47	2.7	1252	10	HS57B2	HS57B2 Homo sapi
122	47	2.7	1364	5	I28326	I28326 Sequence 9	195	47	2.7	71	10	HS57B2	HS57B2 Homo sapi
123	47	2.7	3214	5	I30313	I30313 Sequence 17	196	47	2.7	207957	11	AC004470	AC004470 Homo sapi
124	47	2.7	1393	5	I32738	I32738 Sequence 12	197	47	2.7	133478	11	AC004686	AC004686 Homo sapi
125	47	2.7	399	5	I34395	I34395 Sequence 17	198	47	2.7	586	11	AC005224	AC005224 Homo sapi
126	47	2.7	140	5	I38714	I38714 Sequence 13	199	47	2.7	167525	11	AC005531	AC005531 Homo sapi
127	47	2.7	3214	5	I40334	I40334 Sequence 28	200	47	2.7	2745	11	AF082324	AF082324 Homo sapi
128	47	2.7	120	5	I51760	I51760 Sequence 7	201	47	2.7	2378	11	AF082324	AF082324 Homo sapi
129	47	2.7	966	5	I79595	I79595 Sequence 4	202	47	2.7	1809	11	AF082324	AF082324 Homo sapi
130	47	2.7	120	5	I84386	I84386 Sequence 3	203	47	2.7	4308	11	AF082324	AF082324 Homo sapi
131	47	2.7	120	5	I84386	I84386 Sequence 3	204	47	2.7	4308	11	AF082324	AF082324 Homo sapi
132	47	2.7	120	5	I84386	I84386 Sequence 3	205	47	2.7	4308	11	AF082324	AF082324 Homo sapi
133	47	2.7	120	5	I84386	I84386 Sequence 3	206	47	2.7	4308	11	AF082324	AF082324 Homo sapi
134	47	2.7	120	5	I84386	I84386 Sequence 3	207	47	2.7	4308	11	AF082324	AF082324 Homo sapi
135	47	2.7	120	5	I84386	I84386 Sequence 3	208	47	2.7	4308	11	AF082324	AF082324 Homo sapi
136	47	2.7	120	5	I84386	I84386 Sequence 3	209	47	2.7	4308	11	AF082324	AF082324 Homo sapi
137	47	2.7	120	5	I84386	I84386 Sequence 3	210	47	2.7	4308	11	AF082324	AF082324 Homo sapi
138	47	2.7	120	5	I84386	I84386 Sequence 3	211	47	2.7	4308	11	AF082324	AF082324 Homo sapi
139	47	2.7	120	5	I84386	I84386 Sequence 3	212	47	2.7	4308	11	AF082324	AF082324 Homo sapi
140	47	2.7	120	5	I84386	I84386 Sequence 3	213	47	2.7	4308	11	AF082324	AF082324 Homo sapi
141	47	2.7	120	5	I84386	I84386 Sequence 3	214	47	2.7	4308	11	AF082324	AF082324 Homo sapi
142	47	2.7	120	5	I84386	I84386 Sequence 3	215	47	2.7	4308	11	AF082324	AF082324 Homo sapi
143	47	2.7	120	5	I84386	I84386 Sequence 3	216	47	2.7	4308	11	AF082324	AF082324 Homo sapi
144	47	2.7	120	5	I84386	I84386 Sequence 3	217	47	2.7	4308	11	AF082324	AF082324 Homo sapi
145	47	2.7	120	5	I84386	I84386 Sequence 3	218	47	2.7	4308	11	AF082324	AF082324 Homo sapi
146	47	2.7	120	5	I84386	I84386 Sequence 3	219	47	2.7	4308	11	AF082324	AF082324 Homo sapi
147	47	2.7	120	5	I84386	I84386 Sequence 3	220	47	2.7	4308	11	AF082324	AF082324 Homo sapi

221	47	2.7	2398	12	MMTAF148	Y09972	M.musculus	294	46	2.6	1558	4	XLBMP4	X63426	X.laevis	MR
222	47	2.7	2179	12	MTNFR5	X57796	Mouse mRNA	295	46	2.6	1389	4	XLNLPN038	X56039	X.laevis	MR
223	47	2.7	1858	12	MMTRANS1	X63162	M.musculus	296	46	2.6	2700	4	XLOOCM	X68249	X.laevis	MR
224	47	2.7	5330	12	MMU49908	U49908	Mus musculus	297	46	2.6	3905	5	A03736	A03736	H.sapiens	m
225	47	2.7	4032	12	MMU79523	U79523	Mus musculus	298	46	2.6	80	5	A08907	A08907	H.sapiens	m
226	47	2.7	642	12	MMU95114	U95114	Mus musculus	299	46	2.6	77	5	A08908	A08908	H.sapiens	m
227	47	2.7	2621	12	MUSBRD	L31396	Mus musculus	300	46	2.6	96	5	A08909	A08909	H.sapiens	m
228	47	2.7	3924	12	MUSBRDA	L31397	Mus musculus	301	46	2.6	104	5	A08910	A08910	H.sapiens	m
229	47	2.7	1056	12	MUSUNKNB	L04849	Mouse (clon	302	46	2.6	161	5	A08916	A08916	H.sapiens	m
230	47	2.7	1262	12	MUSUNKNF	L04852	Mouse (clon	303	46	2.6	941	5	A15345	A15345	Bam HI huma	
231	47	2.7	1395	12	RNCA11IC	X66871	R.norvegicu	304	46	2.6	748	5	A21101	A21101	N.tabacum	S
232	47	2.7	3392	12	RNGPHYRN	X60769	Rat sfb mRN	305	46	2.6	347	5	A23327	A23327	O.sativa	mr
233	47	2.7	1194	12	RNSFR	S53987	nicotinic r	306	46	2.6	796	5	A23630	A23630	Corn anthr	
234	47	2.7	3030	12	S53987	S83440	RNH-1=prepr	307	46	2.6	1830	5	A27631	A27631	JGM coat p	
235	47	2.7	2756	12	HSU27131	U27131	Human STS 1	308	46	2.6	1014	5	A30330	A30330	Artificial	
236	47	2.7	492	14	HSU28794	U28794	Human STS 1	309	46	2.6	1016	5	A30331	A30331	Artificial	
237	47	2.7	566	14	HSU28794	U28794	Human STS 1	310	46	2.6	3905	5	A31790	A31790	H.sapiens	m
238	47	2.7	5482	15	AF090435	AF090435	Cloning v	311	46	2.6	1013	5	A32826	A32826	Synthetic c	
239	47	2.7	14225	17	LDVGLYPOL	U76034	Microplitis	312	46	2.6	1015	5	A32827	A32827	Synthetic c	
240	47	2.7	961	17	MDU76034	U76034	Microplitis	313	46	2.6	1837	5	A47357	A47357	Sequence 4	
241	47	2.7	1801	17	TEU27491	U27491	Tick-borne	314	46	2.6	1934	5	A49139	A49139	Sequence 4	
242	47	2.7	11141	17	TEU27491	U27491	Tick-borne	315	46	2.6	3437	5	A52184	A52184	Sequence 1	
243	47	2.7	11141	17	TEU27491	U27491	Tick-borne	316	46	2.6	1347	5	A58523	A58523	Sequence 17	
244	47	2.7	152143	34	AC004932	AC004932	Homo sapi	317	46	2.6	10288	5	A58524	A58524	Sequence 27	
245	47	2.7	113338	34	HSU773A18	AL049557	Homo sapi	318	46	2.6	2400	5	A58545	A58545	Sequence 28	
246	47	2.7	312766	34	PFMALP4	AL035476	Plasmodiu	319	46	2.6	6791	5	A60092	A60092	Sequence 1	
247	47	2.7	182538	34	PFMALP4	AL035477	Plasmodiu	320	46	2.6	10195	5	A60094	A60094	Sequence 2	
248	47	2.7	228625	35	AC004153	AC004153	Plasmodiu	321	46	2.6	1577	5	A65336	A65336	Sequence 4	
249	47	2.7	187544	35	AC004157	AC004157	Plasmodiu	322	46	2.6	1577	5	A65336	A65336	Sequence 59	
250	47	2.7	180278	35	AC007445	AC007445	Homo sapi	323	46	2.6	938	5	A70386	A70386	Sequence 16	
251	47	2.7	160021	35	AC007857	AC007857	Homo sapi	324	46	2.6	1001	5	AR016469	AR016469	Sequence	
252	47	2.7	4875	36	AB009080	AB009080	Dictyoste	325	46	2.6	2267	5	E01314	E01314	CDNA encodl	
253	47	2.7	14955	36	AF000580	AF000580	Dictyoste	326	46	2.6	1924	5	E01356	E01356	CDNA encodl	
254	47	2.7	2253	36	AF032396	AF032396	Hyalophor	327	46	2.6	1861	5	E01357	E01357	CDNA encodl	
255	47	2.7	1743	36	CELRABGDI	U00002	Guanine nuc	328	46	2.6	988	5	E01573	E01573	CDNA sequen	
256	47	2.7	10053	36	CFU19151	U19151	Citithidia f	329	46	2.6	988	5	E01575	E01575	CDNA sequen	
257	47	2.7	678	36	DE5468	X53445	Dictyostei	330	46	2.6	3851	5	E01614	E01614	DNA sequen	
258	47	2.7	681	36	DDE628	X53448	Dictyostei	331	46	2.6	3820	5	E02221	E02221	DNA encodl	
259	47	2.7	1840	36	DDU73686	U73686	Dictyostei	332	46	2.6	958	5	E02319	E02319	DNA encodl	
260	47	2.7	3937	36	DMU35074	U35074	Drosophila	333	46	2.6	1558	5	E03671	E03671	DNA encodl	
261	47	2.7	1432	36	PAU69957	U69957	Periplaneta	334	46	2.6	667	5	E06788	E06788	DNA encodl	
262	47	2.7	108908	36	PFMAL3P8	AL034560	Plasmodiu	335	46	2.6	687	5	E06789	E06789	CDNA encodl	
263	47	2.7	915	36	RPU61144	U61144	Rhodnius pr	336	46	2.6	667	5	E06790	E06790	CDNA encodl	
264	47	2.7	1963	36	SMU8664	U86674	Schistosoma	337	46	2.6	9589	5	E07362	E07362	CDNA of Hep	
265	47	2.7	1288	36	TBBDPMSYN	U86674	Schistosoma	338	46	2.6	826	5	E08516	E08516	DNA encodl	
266	47	2.7	587	36	TOUT89944	U89944	Taenia ovis	339	46	2.6	779	5	E08517	E08517	DNA encodl	
267	47	2.7	3351	37	AF031517	AF031517	Drosophill	340	46	2.6	2430	5	E12586	E12586	CDNA sequen	
268	47	2.7	1104	37	AF047611	AF047611	Euroglyph	341	46	2.6	779	5	E12587	E12587	CDNA encodl	
269	47	2.7	1108	37	AF060449	AF060449	Lima	342	46	2.6	1113	5	E12888	E12888	CDNA encodl	
270	47	2.7	5117	37	AF067198	AF067198	Dictyoste	343	46	2.6	1113	5	E13364	E13364	Plasmid pyg	
271	47	2.7	1412	37	AF083228	AF083228	Caenorhab	344	46	2.6	1769	5	E13524	E13524	3' untransla	
272	47	2.7	1509	37	AF100676	AF100676	Dictyoste	345	46	2.6	622	5	E13525	E13525	3' untransla	
273	47	2.7	3488	37	AF129433	AF129433	Drosophill	346	46	2.6	3851	5	E13540	E13540	Seriola sp.	
274	47	2.7	855	37	AF130840	AF130840	Rhyzopert	347	46	2.6	3451	5	E13998	E13998	Cyanidium c	
275	47	2.7	2856	37	AF132149	AF132149	Drosophill	348	46	2.6	941	5	I03096	I03096	Sequence 8	
276	47	2.7	928	37	AF132151	AF132151	Drosophill	349	46	2.6	688	5	I03321	I03321	Sequence 11	
277	47	2.7	1852	37	AF132556	AF132556	Drosophill	350	46	2.6	1124	5	I06996	I06996	Sequence 12	
278	47	2.7	3143	37	SV3271	SV3271	trehalase l	351	46	2.6	2035	5	I08608	I08608	Sequence 3	
279	47	2.7	1006	37	TVH87097	U87097	Trichomonas	352	46	2.6	1923	5	I09202	I09202	Sequence 1	
280	47	2.7	107402	42	AC007043	AC007043	Homo sapi	353	46	2.6	1423	5	I09360	I09360	Sequence 1	
281	46	2.6	2821	3	AF039137	AF039137	Felis cat	354	46	2.6	1013	5	I11571	I11571	Sequence 23	
282	46	2.6	2815	3	AF039138	AF039138	Felis cat	355	46	2.6	989	5	I12861	I12861	Sequence 2	
283	46	2.6	2805	3	BITACTRA	X57084	Bovine mRNA	356	46	2.6	570	5	I12861	I12861	Sequence 2	
284	46	2.6	2714	3	BITU75304	U75304	Bos laurus	357	46	2.6	570	5	I18355	I18355	Sequence 10	
285	46	2.6	1316	3	OCU131955	AJ131955	Oryctolag	358	46	2.6	1013	5	I21342	I21342	Sequence 10	
286	46	2.6	2836	3	SSU94934	U94934	Sus scrofa	359	46	2.6	2760	5	I24003	I24003	Sequence 30	
287	46	2.6	1665	4	AF080068	AF080068	Xenopus 1	360	46	2.6	2760	5	I25048	I25048	Sequence 3	
288	46	2.6	2085	4	AF080622	AF080622	Danio rer	361	46	2.6	796	5	I33111	I33111	Sequence 3	
289	46	2.6	1030	4	DRHPRABIN	X97332	D. rerio	362	46	2.6	1641	5	I33392	I33392	Sequence 8	
290	46	2.6	1714	4	GDSSDBF	X80340	G. domesticu	363	46	2.6	609	5	I34392	I34392	Sequence 2	
291	46	2.6	1455	4	MSUT18860	Y18860	Mullus surm	364	46	2.6	10136	5	I35495	I35495	Sequence 9	
292	46	2.6	1559	4	OLJ000937	M96857	Oryzias l	365	46	2.6	570	5	I40331	I40331	Sequence 2	
293	46	2.6	2262	4	XELCDC25AA	M96857	Xenopus lae	366	46	2.6	1817	5	I42402	I42402	Sequence 5	

367	46	2.6	748	5	I44072	144072	Sequence 3	C	440	46	2.6	107135	9	HS145B12	AL008706 Human DNA
368	46	2.6	1898	5	I46765	146765	Sequence 1		441	46	2.6	36429	9	HS14B7	249258 Human DNA s
369	46	2.6	347	5	I47730	147730	Sequence 2		442	46	2.6	103144	9	HS154K9	294237 Human DNA s
370	46	2.6	140	5	I48978	148978	Sequence 5		443	46	2.6	103146	9	HS187N21	268036 Human DNA s
371	46	2.6	240	5	I48879	148879	Sequence 6		444	46	2.6	120376	9	HS271G9	292543 Human DNA s
372	46	2.6	289	5	I63077	163077	Sequence 3		445	46	2.6	143747	9	HS370M22	282206 Human DNA s
373	46	2.6	1558	5	I66401	166401	Sequence 7		446	46	2.6	170952	9	HS501N12	AL022170 Homo sapi
374	46	2.6	5173	5	I68732	168732	Sequence 1		447	46	2.6	176433	9	HS711L6	AL022165 Human DNA
375	46	2.6	2405	5	I73428	173428	Sequence 30		448	46	2.6	3723	9	HSADTG	Y12226 H. sapiens m
376	46	2.6	216	5	I80062	180062	Sequence 34		449	46	2.6	2000	9	HSXO28ORF	X99270 H. sapiens X
377	46	2.6	69142	7	AB012239	AB012239	Arabidops		450	46	2.6	4280	9	H0M130LEU	M92349 Human leuci
378	46	2.6	2440	7	AB021862	AB021862	Cucumis m		451	46	2.6	2287	9	H0M42PIA	D00174 Homo sapien
379	46	2.6	1206	7	ATANNEX	X99224 A. thaliana			452	46	2.6	2271	9	H0M4SP	J05032 Human aspar
380	46	2.6	513	7	ATCOR66M	X55053 A. thaliana			453	46	2.6	1266	9	H0MDP2M	L40386 Human DP--2
381	46	2.6	835	7	ATP27A	Y11792 A. thaliana			454	46	2.6	1659	9	H0MP57B	D44497 Human trans
382	46	2.6	680	7	ATSRP14	Y10116 A. thaliana			455	46	2.6	6075	9	H0MTN12	M80340 Human trans
383	46	2.6	1770	7	BNPML2	X73849 B. napus (pn			456	46	2.6	2549	9	U00686	U00686 Human putat
384	46	2.6	1104	7	BOMIPATCP	X93639 B. oleracea			457	46	2.6	134218	10	HS170F5	AL035067 Human DNA
385	46	2.6	1419	7	CCCHSMR	Z67988 C. chinensis			458	46	2.6	117338	10	HS173D1	AL031984 Human DNA
386	46	2.6	783	7	CCFCPI0	AJ000972 Cyclolell			459	46	2.6	69017	10	HS410I8	AL0341732 Human DNA
387	46	2.6	987	7	CLJACOIV	X57832 C. lacryma-1			460	46	2.6	101270	10	HS483K16	AL034374 Human DNA
388	46	2.6	673	7	CRPS27	X83694 C. reinhardt			461	46	2.6	130467	10	HS798A17	AL031274 Human DNA
389	46	2.6	1839	7	D14059	D14059 Ananas como			462	46	2.6	1608	10	HSDA23	X83544 H. sapiens D
390	46	2.6	1735	7	DCRCBT3	Z84571 D. carophytl			463	46	2.6	39015	10	HSE90C2	252182 Homo sapien
391	46	2.6	2263	7	HVRNAGAM1	X87690 H. vulgare m			464	46	2.6	5959	10	HSB8M	X72889 H. sapiens h
392	46	2.6	2641	7	LENLEPRIB	D14489 Lentinus ed			465	46	2.6	1817	10	HSICAM3RN	X69819 H. sapiens I
393	46	2.6	1236	7	LOLOPIB	M59163 Lolium perit			466	46	2.6	5994	10	HSINTB4R	X53587 Human mRNA
394	46	2.6	1810	7	MCPTKINA	Z30329 M. crystalli			467	46	2.6	3688	10	HSKCHMR	X65877 H. sapiens m
395	46	2.6	1162	7	MNHRBCSA	M95993 Manihot esc			468	46	2.6	1483	10	HSM800068	AL049283 Homo sapi
396	46	2.6	837	7	NPSODM	X55974 N. plumbagin			469	46	2.6	1772	10	HSM800085	AL049300 Homo sapi
397	46	2.6	2400	7	NPZEAXANT	X95732 N. plumbagin			470	46	2.6	580	10	HSM800144	AL050277 Homo sapi
398	46	2.6	2851	7	OSGOG1	Y12594 O. sativa mr			471	46	2.6	3386	10	HSM800161	AL050092 Homo sapi
399	46	2.6	4120	7	OSGOG2	Y12595 O. sativa mr			472	46	2.6	1694	10	HSM800167	AL049382 Homo sapi
400	46	2.6	875	7	OSRTA2PR	AJ001265 Oryza sat			473	46	2.6	669	10	HSM800251	AL049466 Homo sapi
401	46	2.6	1391	7	PCU10306	U10306 Phanerochae			474	46	2.6	1438	10	HSM800279	AL049838 Homo sapi
402	46	2.6	2159	7	PSDPS	Y08611 P. sativum m			475	46	2.6	2785	10	HSM800418	AL050108 Homo sapi
403	46	2.6	575	7	PSGAMRA	X65154 P. sativum g			476	46	2.6	2462	10	HSM800419	AL050116 Homo sapi
404	46	2.6	708	7	PSGDCPG	X54377 P. sativum g			477	46	2.6	3469	10	HSM800438	AL050143 Homo sapi
405	46	2.6	518	7	QSAJ692	AJ000692 Quercus s			478	46	2.6	2765	10	HSM800443	AL050143 Homo sapi
406	46	2.6	498	7	S47413	S47413 glycine-ric			479	46	2.6	959	10	HSMPZE	X61970 H. sapiens m
407	46	2.6	900	7	SLDEHYDLA	Z21500 S. longipes			480	46	2.6	380	10	HSMUC5BR	Y10080 H. sapiens m
408	46	2.6	2235	7	SOPMOPN	Z66599 S. oleracea			481	46	2.6	2200	10	HSMXR7MR	Z37987 H. sapiens m
409	46	2.6	3437	7	SOPULSPD	X83969 S. oleracea			482	46	2.6	1550	10	HSMXR9	Z49216 H. sapiens m
410	46	2.6	795	7	SOVSCAM1X	L01430 Soybean cal			483	46	2.6	39000	10	HSN86D4	Z82250 Human DNA s
411	46	2.6	1432	7	SPCIG2	X70046 S. pombe mrn			484	46	2.6	2739	10	HSTRANSLI	X78627 H. sapiens m
412	46	2.6	2062	7	WHTEPI1X	M90077 Wheat trans			485	46	2.6	1368	10	HSY16645	Y16645 Homo sapien
413	46	2.6	1016	8	AF005158	AF005158 Arabidops			486	46	2.6	2123	10	S77771S1	S77772 Ku autoanti
414	46	2.6	610	8	AF009959	AF009959 Oryza sat			487	46	2.6	837	10	S78214	S78214 APC-tumor s
415	46	2.6	1592	8	AF019630	AF019630 Magnapor			488	46	2.6	125590	11	AC002382	AC002382 Human BAC
416	46	2.6	764	8	AF022731	AF022731 Oryza sat			489	46	2.6	128915	11	AC002416	AC002416 Human BAC
417	46	2.6	1032	8	AF043538	AF043538 Arabidops			490	46	2.6	128915	11	AC002464	AC002464 Human BAC
418	46	2.6	747	8	AF055372	AF055372 Arabidops			491	46	2.6	123805	11	AC002500	AC002500 Human Cos
419	46	2.6	1663	8	AF058763	AF058763 zea may			492	46	2.6	41084	11	AC002531	AC002531 Homo sapi
420	46	2.6	2351	8	AF060518	AF060518 Cuphea pu			493	46	2.6	197800	11	AC002540	AC002540 Homo sapi
421	46	2.6	2323	8	AF071888	AF071888 Prunus ar			494	46	2.6	70851	11	AC002564	AC002564 Human BAC
422	46	2.6	1700	8	AF121261	AF121261 Lilium lo			495	46	2.6	94320	11	AC003005	AC003005 Human DNA
423	46	2.6	1647	8	AF124842	AF124842 Capsicum			496	46	2.6	45084	11	AC003042	AC003042 Homo sapi
424	46	2.6	1404	8	ATU62741	U62741 Arabidops			497	46	2.6	102818	11	AC003046	AC003046 Homo sapi
425	46	2.6	1492	8	H4U91341	U91341 Helianthus			498	46	2.6	113457	11	AC003685	AC003685 Homo sapi
426	46	2.6	801	8	RICLMMHSP	M80186 Oryza sativ			499	46	2.6	154252	11	AC004062	AC004062 Homo sapi
427	46	2.6	1276	8	S59422	S59422 Populus x c			500	46	2.6	42007	11	AC004200	AC004200 Homo sapi
428	46	2.6	826	8	S65048	S65048 self-incorp			501	46	2.6	156461	11	AC004227	AC004227 Homo sapi
429	46	2.6	1425	8	SOYSPK3	L19361 Glycine max			502	46	2.6	185516	11	AC004383	AC004383 Human Chr
430	46	2.6	1154	8	TAU87163	U87163 Triticum ae			503	46	2.6	80659	11	AC004485	AC004485 Homo sapi
431	46	2.6	3901	8	VUA225088	AJ225088 Vigna ung			504	46	2.6	127824	11	AC004553	AC004553 Homo sapi
432	46	2.6	45966	9	AP000028	AP000028 Homo sapi			505	46	2.6	127824	11	AC004822	AC004822 Homo sapi
433	46	2.6	100000	9	AP000130	AP000130 Homo sapi			506	46	2.6	85058	11	AC004989	AC004989 Homo sapi
434	46	2.6	100000	9	AP000208	AP000208 Homo sapi			507	46	2.6	73899	11	AC005156	AC005156 Homo sapi
435	46	2.6	9918	9	AP000218	AP000218 Homo sapi			508	46	2.6	198582	11	AC005291	AC005291 Homo sapi
436	46	2.6	115118	9	AP000247	AP000247 Homo sapi			509	46	2.6	131070	11	AC005295	AC005295 Homo sapi
437	46	2.6	35932	9	AP000340	AP000340 Homo sapi			510	46	2.6	87114	11	AC005353	AC005353 Homo sapi
438	46	2.6	1681	9	D89079	D89079 Homo sapien			511	46	2.6	16167	11	AC005483	AC005483 Homo sapi
439	46	2.6	1216	9	EN1838	AJ001838 Homo sapi			512	46	2.6	183085	11	AC005815	AC005815 complet

C 513	513	4/6	2.6	134045	11	AC00586	Homo sapi	586	4/6	2.6	417	17	TBEV3UT	Y08863	Tick-borne
C 514	514	4/6	2.6	178367	11	AC006115	Homo sapi	587	4/6	2.6	768	17	TEU27490	U073401	Tick-borne
C 515	515	4/6	2.6	154354	11	AC006197	Homo sapi	C 588	4/6	2.6	9589	26	E07361	E07361	gRNA of Hep
C 516	516	4/6	2.6	79085	11	AC006221	Homo sapi	C 589	4/6	2.6	1657	26	E11332	E11332	Human cDNA
C 517	517	4/6	2.6	220218	11	AC006288	Homo sapi	C 590	4/6	2.6	32208	34	AC004105	AC004105	Homo sapi
C 518	518	4/6	2.6	186555	11	AC006313	Homo sapi	C 591	4/6	2.6	206056	34	AC004387	AC004387	Homo sapi
C 519	519	4/6	2.6	1728	11	AF013349	Homo sapi	C 592	4/6	2.6	170000	34	AC004578	AC004578	Homo sapi
C 520	520	4/6	2.6	2150	11	AF017790	Homo sapi	C 593	4/6	2.6	190000	34	AC004580	AC004580	Homo sapi
C 521	521	4/6	2.6	1572	11	AF035305	Homo sapi	C 594	4/6	2.6	188477	34	AC004971	AC004971	Homo sapi
C 522	522	4/6	2.6	2632	11	AF040751	Homo sapi	C 595	4/6	2.6	171188	34	AC005025	AC005025	Homo sapi
C 523	523	4/6	2.6	2843	11	AF044323	Homo sapi	C 596	4/6	2.6	180593	34	AC005282	AC005282	Homo sapi
C 524	524	4/6	2.6	2098	11	AF060866	Homo sapi	C 597	4/6	2.6	146209	34	AC005522	AC005522	Homo sapi
C 525	525	4/6	2.6	2619	11	AF068829	Homo sapi	C 598	4/6	2.6	41755	34	AC006178	AC006178	Homo sapi
C 526	526	4/6	2.6	5075	11	AF081195	Homo sapi	C 599	4/6	2.6	195764	34	AC006483	AC006483	Homo sapi
C 527	527	4/6	2.6	2802	11	AF084644	Homo sapi	C 600	4/6	2.6	124673	34	HS2394	HS2394	Homo sapi
C 528	528	4/6	2.6	2905	11	AF084645	Homo sapi	C 601	4/6	2.6	112555	34	HS419C19	HS419C19	Homo sapi
C 529	529	4/6	2.6	1212	11	AF100781	Homo sapi	C 602	4/6	2.6	115272	34	HS475N6C	HS475N6C	Homo sapi
C 530	530	4/6	2.6	56887	11	HSMHCC345	Homo sapien	C 603	4/6	2.6	91029	34	HS72E17	HS72E17	Homo sapi
C 531	531	4/6	2.6	2902	11	HSU31501	Homo sapien	C 604	4/6	2.6	133505	34	HS785G19	HS785G19	Homo sapi
C 532	532	4/6	2.6	2516	11	HSU37359	Homo sapien	C 605	4/6	2.6	169254	34	HS892P23	HS892P23	Homo sapi
C 533	533	4/6	2.6	5173	11	HSU38847	Homo sapien	C 606	4/6	2.6	142807	34	HSBJ02553	HSBJ02553	Homo sapi
C 534	534	4/6	2.6	1280	11	HSU54559	Homo sapien	C 607	4/6	2.6	110000	34	HSDB1689N3_1	HSDB1689N3_1	Continuation (2
C 535	535	4/6	2.6	2814	11	HSU73682	Homo sapien	C 608	4/6	2.6	263628	34	HSDB784K2	HSDB784K2	Homo sapi
C 536	536	4/6	2.6	42179	11	HSU76377	Human menin	C 609	4/6	2.6	247159	34	HSDB81F6	HSDB81F6	Homo sapi
C 537	537	4/6	2.6	8745	11	HSU891329	Human olfac	C 610	4/6	2.6	123100	34	HSJ72T10	HSJ72T10	Homo sapi
C 538	538	4/6	2.6	4082	11	HSU91836	Human prote	C 611	4/6	2.6	111165	34	HSJ842K24	HSJ842K24	Homo sapi
C 539	539	4/6	2.6	69015	11	HSU95739	Human kines	C 612	4/6	2.6	300172	35	AC005308	AC005308	Homo sapi
C 540	540	4/6	2.6	188636	11	HUAC002287	Human chrom	C 613	4/6	2.6	100932	35	AC005504	AC005504	Homo sapi
C 541	541	4/6	2.6	130491	11	HUAC003977	Homo sapi	C 614	4/6	2.6	258033	35	AC005507	AC005507	Plasmid
C 542	542	4/6	2.6	3828	12	AC004399	Mus muscu	C 615	4/6	2.6	155771	35	AC005902	AC005902	Homo sapi
C 543	543	4/6	2.6	4128	12	AF022363	Mus muscu	C 616	4/6	2.6	169387	35	AC006338	AC006338	Homo sapi
C 544	544	4/6	2.6	2246	12	AF026122	Mus muscu	C 617	4/6	2.6	205380	35	AC006373	AC006373	Homo sapi
C 545	545	4/6	2.6	10919	12	AF031572	Mus muscu	C 618	4/6	2.6	199600	35	AC006460	AC006460	Homo sapi
C 546	546	4/6	2.6	4452	12	AF054831	Mus muscu	C 619	4/6	2.6	122459	35	AC006944	AC006944	Mus muscu
C 547	547	4/6	2.6	3895	12	AF079765	Mus muscu	C 620	4/6	2.6	122459	35	AC006944	AC006944	Mus muscu
C 548	548	4/6	2.6	1661	12	AF081366	Rattus no	C 621	4/6	2.6	190653	35	AC006994	AC006994	Homo sapi
C 549	549	4/6	2.6	1315	12	AF082526	Mus muscu	C 622	4/6	2.6	195550	35	AC006995	AC006995	Homo sapi
C 550	550	4/6	2.6	965	12	AF106945	Rattus no	C 623	4/6	2.6	171771	35	AC007038	AC007038	Homo sapi
C 551	551	4/6	2.6	4365	12	AF114168	Mus muscu	C 624	4/6	2.6	182480	35	AC007056	AC007056	Homo sapi
C 552	552	4/6	2.6	834	12	AF115410	Cricetulu	C 625	4/6	2.6	168575	35	AC007078	AC007078	Homo sapi
C 553	553	4/6	2.6	1856	12	AF143957	Mus muscu	C 626	4/6	2.6	193800	35	AC007391	AC007391	Homo sapi
C 554	554	4/6	2.6	154614	12	MM573K1	Mus muscu	C 627	4/6	2.6	193301	35	AC007392	AC007392	Homo sapi
C 555	555	4/6	2.6	251364	12	MMAE000654	Mus muscu	C 628	4/6	2.6	187817	35	AC007397	AC007397	Homo sapi
C 556	556	4/6	2.6	158405	12	MMHC0310M6	Mus muscu	C 629	4/6	2.6	167373	35	AC007604	AC007604	Homo sapi
C 557	557	4/6	2.6	2997	12	MMPC1A13A	Mus muscu	C 630	4/6	2.6	328876	35	AC007610	AC007610	Homo sapi
C 558	558	4/6	2.6	1778	12	MMU02885	Mus muscu	C 631	4/6	2.6	191300	35	AC007749	AC007749	Homo sapi
C 559	559	4/6	2.6	3715	12	MMU35846	Mus muscu	C 632	4/6	2.6	925	36	AEVA0440X	AEVA0440X	Homo sapi
C 560	560	4/6	2.6	2693	12	MMU83172	Mus muscu	C 633	4/6	2.6	4028	36	AF003306	AF003306	Drosophila
C 561	561	4/6	2.6	756	12	MRBPRL7B	M. musculus	C 634	4/6	2.6	2062	36	AF035482	AF035482	Homo sapi
C 562	562	4/6	2.6	901	12	MUSMASC12	Mus muscu	C 635	4/6	2.6	40753	36	BEY18367	BEY18367	Branchiost
C 563	563	4/6	2.6	795	12	MUSUNNL	Mus muscu	C 636	4/6	2.6	1355	36	CEU68266	CEU68266	Caenorhabdi
C 564	564	4/6	2.6	814	12	RATPHGP	Rattus norv	C 637	4/6	2.6	763	36	CEU96172	C6172	C. tritid
C 565	565	4/6	2.6	1461	12	RNMATFA	R. norvegicu	C 638	4/6	2.6	1907	36	CVA00A0R	CVA00A0R	C. viridis
C 566	566	4/6	2.6	409	12	RNMTCYTOX	X06146	C 639	4/6	2.6	479	36	DDCOX6R	DDCOX6R	D. discoidei
C 567	567	4/6	2.6	1837	12	RNP2XMR	X80477	C 640	4/6	2.6	6428	36	DDPRE	DDPRE	D. discoidei
C 568	568	4/6	2.6	2153	12	RNP4502C	Rat mRNA fo	C 641	4/6	2.6	91	36	DDIACCTN8A	DDIACCTN8A	D. discoidei
C 569	569	4/6	2.6	2225	12	RNU30290	Rattus norv	C 642	4/6	2.6	2581	36	DDIGP138A	DDIGP138A	D. discoidei
C 570	570	4/6	2.6	1560	12	RNU76419	Rattus norv	C 643	4/6	2.6	1128	36	DDIPORIN	DDIPORIN	D. discoidei
C 571	571	4/6	2.6	1837	12	S69385	epithelial	C 644	4/6	2.6	1946	36	DDIRRS	DDIRRS	D. discoidei
C 572	572	4/6	2.6	2804	12	S69407	Edg-endogli	C 645	4/6	2.6	2415	36	DDPREDE	DDPREDE	D. discoidei
C 573	573	4/6	2.6	3180	12	S79832	transcripti	C 646	4/6	2.6	5890	36	DDU60086	DDU60086	D. discoidei
C 574	574	4/6	2.6	1165	12	S83456	GTP-binding	C 647	4/6	2.6	2441	36	DDU57940	DDU57940	D. discoidei
C 575	575	4/6	2.6	371	14	AMU73923	Apis mellif	C 648	4/6	2.6	3504	36	DMKRN1	DMKRN1	Drosophila
C 576	576	4/6	2.6	432	14	G36768	SHGC-5482	C 649	4/6	2.6	140	36	DMKRN1	DMKRN1	Drosophila
C 577	577	4/6	2.6	8414	15	CVU30496	Cloning vec	C 650	4/6	2.6	1428	36	DMRS55MR	DMRS55MR	Drosophila
C 578	578	4/6	2.6	8392	15	CVU30497	Cloning vec	C 651	4/6	2.6	1376	36	DMSCPXRTH	DMSCPXRTH	D. melanogass
C 579	579	4/6	2.6	791	16	OSA004966	Oryza sat	C 652	4/6	2.6	1482	36	HERFAMDE	HERFAMDE	H. echinatis
C 580	580	4/6	2.6	214	17	D67092	Hepatitlis C	C 653	4/6	2.6	691	36	HVI48594	HVI48594	Heliothia
C 581	581	4/6	2.6	1208	17	HPCHC15T70	Hepatitlis C	C 654	4/6	2.6	1601	36	LPACTS	LPACTS	L. polyphenu
C 582	582	4/6	2.6	1227	17	HPCHNSJ7	Hepatitlis C	C 655	4/6	2.6	1567	36	LTRCT1567	LTRCT1567	L. ferestrii
C 583	583	4/6	2.6	9289	17	HPCPOLP	Hepatitlis C	C 656	4/6	2.6	1603	36	LTRCT1503	LTRCT1503	L. ferestrii
C 584	584	4/6	2.6	9456	17	HPCRN1	Hepatitlis C	C 657	4/6	2.6	1577	36	PELCP4K	PELCP4K	Periplaneta
C 585	585	4/6	2.6	12226	17	RTUPOLYP	Rice tungro	C 658	4/6	2.6	2633	36	PFASPHM	PFASPHM	F. talciparu

659	46	2.6	3528	36	PF016995	U16995 Plasmodium	732	45	2.6	2123	4	GGU77351	U77351 Gallus gall
C 660	46	2.6	1538	36	PFU97128	U97128 Plasmodium	733	45	2.6	1620	4	GGVITRO	Y11030 G.gallus mr
C 661	46	2.6	476	36	PY18125	Y18125 Physarum po	734	45	2.6	1170	4	HNPROLAC	X61049 H.nobilis m
662	46	2.6	869	36	S75286	S75286 major aller	735	45	2.6	1544	4	OLANEXM3	Y11254 O.laticpes m
663	46	2.6	613	36	S76337	S76337 tge-binding	736	45	2.6	2728	4	OMUJ3118	AO00318 Oncorhnc
664	46	2.6	1059	36	S76894	S76894 Daep-1-cuti	737	45	2.6	586	4	OMU83980	U83980 Oncorhnc
C 665	46	2.6	2532	36	SLMAGTARD	M15272 slime mold	738	45	2.6	443	4	PBDESRPB	X72387 P.bicolor m
666	46	2.6	1407	36	SPMUSACT	X61042 S.plicata m	739	45	2.6	958	4	S54890	S54890 growth horm
668	46	2.6	7291	36	TRBRETE	Y13395 Trypanosoma	740	45	2.6	994	4	S56212	S56212 nerve growt
669	46	2.6	953	36	TVCP4	X77221 T.vaginalis	741	45	2.6	1934	4	S59519	S59519 stannocalc
C 670	46	2.6	13674	36	U00691	U00691 Dictyostell	742	45	2.6	5078	4	TNU75604	U75604 Tiliapia nil
C 671	46	2.6	13729	36	U00796	U00796 Dictyostell	743	45	2.6	542	4	XEHIS31A	XEHIS31A xenopus lae
C 672	46	2.6	12530	37	AE001417	AE001417 Plasmodi	744	45	2.6	2526	4	XHCFIN	XHCFIN xipophorus
673	46	2.6	2435	37	AF007768	AF007768 Choriston	745	45	2.6	1469	4	XLB506AP4	X87582 X.laeyis B-
674	46	2.6	790	37	AF040384	AF040384 Schistos	746	45	2.6	2134	4	XLCTGT	Y11587 X.laeyis Cc
675	46	2.6	845	37	AF047612	AF047612 Euroglyph	747	45	2.6	1500	4	XLNKATP	Y11587 X.laeyis mr
676	46	2.6	1429	37	AF072219	AF072219 Blatella	748	45	2.6	2729	4	XLU68387	U68387 Xenopus lae
677	46	2.6	2304	37	AF080675	AF080675 Dictyoste	749	45	2.6	1417	4	ZOPX18678	Y16678 Zosterisess
678	46	2.6	905	37	AF081800	AF081800 Dictyoste	750	45	2.6	1394	5	A07588	A07588 Synthetic H
679	46	2.6	1412	37	AF083227	AF083227 Caenorhab	751	45	2.6	2399	5	A07647	A07647 Synthetic g
680	46	2.6	477	37	AF098065	AF098065 Trypanos	752	45	2.6	3347	5	A12297	A12297 Rit beta cd
681	46	2.6	6396	37	AF102268	AF102268 Pacifasta	753	45	2.6	7859	5	A17115	A17115 yeast expe
682	46	2.6	3754	37	AF132141	AF132141 Drosophil	754	45	2.6	80	5	A18079	A18079 yeast expe
683	46	2.6	2330	37	AF132142	AF132142 Drosophil	755	45	2.6	11517	5	A18777	A18777 5' terminus
684	46	2.6	2008	37	AF145689	AF145689 Drosophil	756	45	2.6	7859	5	A18788	A18788 complete nu
685	46	2.6	446	37	AF146743	AF146743 Mesobuthu	757	45	2.6	1859	5	A19596	A19596 SEQ ID NO:
686	46	2.6	1636	37	CEU60058	U92647 Geocarcinus	758	45	2.6	1284	5	A20553	A20553 NF-YB. 6/19
687	46	2.6	1459	37	GIU92647	U92647 Geocarcinus	759	45	2.6	1046	5	A21103	A21103 N.tabacum S
688	46	2.6	1626	37	OUV81605	U81605 Onchocerca	760	45	2.6	3138	5	A21625	A21625 expression
689	46	2.6	2846	37	TOXATPC	L39079 Toxoplasma	761	45	2.6	2966	5	A27171	A27171 Tyrosine Ki
C 690	46	2.6	40933	42	AC002471	AC002471 Homo sapi	762	45	2.6	1966	5	A31001	A31001 G.gallus DN
691	46	2.6	41958	42	AC005374	AC005374 Homo sapi	763	45	2.6	1395	5	A41575	A41575 Sequence 1
692	46	2.6	176257	42	AC006039	AC006039 Homo sapi	764	45	2.6	2594	5	A45787	A45787 Sequence 25
693	46	2.6	191145	42	AC006112	AC006112 Homo sapi	765	45	2.6	2643	5	A47363	A47363 Sequence 10
C 694	46	2.6	126732	42	AC007360	AC007360 Homo sapi	766	45	2.6	141	5	A51774	A51774 Sequence 9
695	46	2.6	80858	42	AF081571	AF081571 Homo sapi	767	45	2.6	3200	5	A52563	A52563 Sequence 1
C 696	46	2.6	3016	42	AF104032	AF104032 Homo sapi	768	45	2.6	1376	5	A57389	A57389 Sequence 4
697	46	2.6	4670	42	AF106697	AF106697 Homo sapi	769	45	2.6	1713	5	A59344	A59344 Sequence 1
698	46	2.6	2180	42	HSU58996	U58996 Homo sapien	770	45	2.6	1441	5	A59403	A59403 Sequence 63
699	46	2.6	1853	42	AF010191	AF010191 Felis cat	771	45	2.6	1141	5	A65340	A65340 Sequence 3
700	46	2.6	3227	3	AF019298	AF019298 Bos tauru	772	45	2.6	1040	5	A65943	A65943 Sequence 3
701	46	2.6	4211	3	AF026008	AF026008 Oryctolag	773	45	2.6	1713	5	A65965	A65965 Sequence 25
702	46	2.6	2494	3	AF126247	AF126247 Canis fam	774	45	2.6	1602	5	AR000496	AR000496 Sequence
703	46	2.6	4697	3	BTTHROW2	X96540 B.taurus mr	775	45	2.6	11517	5	AR001293	AR001293 Sequence
704	46	2.6	1024	3	BTU79414	X79414 Bos taurus	776	45	2.6	80	5	AR005011	AR005011 Sequence
705	46	2.6	5425	3	CEP180RC	X87224 Canis famli	777	45	2.6	1425	5	AR005195	AR005195 Sequence
706	46	2.6	3334	3	CEFSRE72	X67813 C.familiari	778	45	2.6	6671	5	AR011880	AR011880 Sequence
707	46	2.6	1065	3	CHAF000167	AF000167 Capra hir	779	45	2.6	1181	5	AR016802	AR016802 Sequence
708	46	2.6	442	3	OCU58653	U58653 Oryctolagus	780	45	2.6	1147	5	AR020828	AR020828 Sequence
709	46	2.6	763	3	PICAC3A	M84133 Sus scrofa	781	45	2.6	2656	5	E00617	E00617 CDNA encodi
710	46	2.6	462	3	RABMHDP2	M22640 Rabbid MHC	782	45	2.6	2656	5	E00717	E00717 CDNA encodi
711	46	2.6	2763	3	SSAPDAI	X59414 Porcine Apo	783	45	2.6	2235	5	E00778	E00778 Rat liver c
712	46	2.6	1573	3	SSCKRIB	X56503 Porcine CXI	784	45	2.6	669	5	E00984	E00984 Probe detec
713	46	2.6	776	3	SSU37312	U90884 Sus scrofa	785	45	2.6	1973	5	E01812	E01812 CDNA encodi
714	46	2.6	2588	3	SSU90884	AF029404 Xenopus 1	786	45	2.6	669	5	E01963	E01963 CDNA encodi
715	46	2.6	1744	3	SSU92404	AF053120 Xenopus 1	787	45	2.6	1014	5	E02152	E02152 DNA encodin
716	46	2.6	1339	3	AF053120	AF053120 Xenopus 1	788	45	2.6	3630	5	E02914	E02914 CDNA encodi
717	46	2.6	1250	3	AF061981	AF061981 Xenopus 1	789	45	2.6	3429	5	E03202	E03202 DNA encodin
718	46	2.6	2896	3	AF065135	AF065135 Xenopus 1	790	45	2.6	1259	5	E03348	E03348 CDNA sequen
719	46	2.6	5138	3	AF090432	AF090432 Dantio rer	791	45	2.6	1581	5	E03349	E03349 CDNA sequen
720	46	2.6	1770	4	CCZP2A	X27491 C.carpio mr	792	45	2.6	2662	5	E04233	E04233 DNA encodin
721	46	2.6	1004	4	CHKGR78	M27260 Chicken 78-	793	45	2.6	1493	5	E04257	E04257 DNA encodin
722	46	2.6	4577	4	DRAPRIPAT	U89295 Dantio rerio	794	45	2.6	2339	5	E07108	E07108 CDNA encodi
723	46	2.6	1111	4	GNKX28	Y10655 G.gallus NK	795	45	2.6	1632	5	E07941	E07941 CDNA encodi
724	46	2.6	1111	4	GNKX28	Y10655 G.gallus NK	796	45	2.6	9487	5	E08264	E08264 CDNA of Hep
725	46	2.6	1111	4	GNKX28	Y10655 G.gallus NK	797	45	2.6	9487	5	E08443	E08443 DNA fragmen
726	46	2.6	1111	4	GNKX28	Y10655 G.gallus NK	798	45	2.6	9487	5	E08443	E08443 DNA fragmen
727	46	2.6	1111	4	GNKX28	Y10655 G.gallus NK	799	45	2.6	9487	5	E08443	E08443 DNA fragmen
728	46	2.6	1111	4	GNKX28	Y10655 G.gallus NK	800	45	2.6	9487	5	E08443	E08443 DNA fragmen
729	46	2.6	1111	4	GNKX28	Y10655 G.gallus NK	801	45	2.6	9487	5	E08443	E08443 DNA fragmen
730	46	2.6	1111	4	GNKX28	Y10655 G.gallus NK	802	45	2.6	9487	5	E08443	E08443 DNA fragmen
731	46	2.6	1111	4	GNKX28	Y10655 G.gallus NK	803	45	2.6	9487	5	E08443	E08443 DNA fragmen
732	46	2.6	1111	4	GNKX28	Y10655 G.gallus NK	804	45	2.6	9487	5	E08443	E08443 DNA fragmen
733	46	2.6	1111	4	GNKX28	Y10655 G.gallus NK	805	45	2.6	9487	5	E08443	E08443 DNA fragmen
734	46	2.6	1111	4	GNKX28	Y10655 G.gallus NK	806	45	2.6	9487	5	E08443	E08443 DNA fragmen
735	46	2.6	1111	4	GNKX28	Y10655 G.gallus NK	807	45	2.6	9487	5	E08443	E08443 DNA fragmen
736	46	2.6	1111	4	GNKX28	Y10655 G.gallus NK	808	45	2.6	9487	5	E08443	E08443 DNA fragmen
737	46	2.6	1111	4	GNKX28	Y10655 G.gallus NK	809	45	2.6	9487	5	E08443	E08443 DNA fragmen
738	46	2.6	1111	4	GNKX28	Y10655 G.gallus NK	810	45	2.6	9487	5	E08443	E08443 DNA fragmen
739	46	2.6	1111	4	GNKX28	Y10655 G.gallus NK	811	45	2.6	9487	5	E08443	E08443 DNA fragmen
740	46	2.6	1111	4	GNKX28	Y10655 G.gallus NK	812	45	2.6	9487	5	E08443	E08443 DNA fragmen
741	46	2.6	1111	4	GNKX28	Y10655 G.gallus NK	813	45	2.6	9487	5	E08443	E08443 DNA fragmen
742	46	2.6	1111	4	GNKX28	Y10655 G.gallus NK	814	45	2.6	9487	5	E08443	E08443 DNA fragmen
743	46	2.6	1111	4	GNKX28	Y10655 G.gallus NK	815	45	2.6	9487	5	E08443	E08443 DNA fragmen
744	46	2.6	1111	4	GNKX28	Y10655 G.gallus NK	816	45	2.6	9487	5	E08443	E08443 DNA fragmen
745	46	2.6	1111	4	GNKX28	Y10655 G.gallus NK	817	45	2.6	9487	5	E08443	E08443 DNA fragmen
746	46	2.6	1111	4	GNKX28	Y10655 G.gallus NK	818	45	2.6	9487	5	E08443	E08443 DNA fragmen
747	46	2.6	1111	4	GNKX28	Y10655 G.gallus NK	819	45	2.6	9487	5	E08443	E08443 DNA fragmen
748	46	2.6	1111	4	GNKX28	Y10655 G.gallus NK	820	45	2.6	9487	5	E08443	E08443 DNA fragmen
749	46	2.6	1111	4	GNKX28	Y10655 G.gallus NK	821	45	2.6	9487	5	E08443	E08443 DNA fragmen
750	46	2.6	1111	4	GNKX28	Y10655 G.gallus NK	822	45	2.6	9487	5	E08443	E08443 DNA fragmen
751	46	2.6	1111	4	GNKX28	Y10655 G.gallus NK	823	45	2.6	9487	5	E08443	E08443 DNA fragmen
752	46	2.6	1111	4	GNKX28	Y10655 G.gallus NK	824	45	2.6	9487	5	E08443	E08443 DNA fragmen
753	46	2.6	1111	4	GNKX28	Y10655 G.gallus NK	825	45	2.6	9487	5	E08443	E08443 DNA fragmen

C 805	45	2.6	90	5	E12579	E12579	Probe. 6/19	878	45	2.6	899	7	BYNCALMOD	L14071 Bryonia dio
C 805	45	2.6	90	5	E12580	E12580	Probe. 6/19	879	45	2.6	742	7	CACORAT6	X71000 C. arellana
C 807	45	2.6	189	5	E13052	E13052	CDNA encodi	880	45	2.6	1261	7	CAR012751	AJ012751 Cicer ari
C 808	45	2.6	90	5	E13639	E13639	Probe. 6/19	881	45	2.6	2795	7	CBKPMETGN	Z49150 C. blumei K1
C 809	45	2.6	896	5	I00734	I00734	Sequence 3	882	45	2.6	630	7	CHH111	X77253 C. herbarum
C 810	45	2.6	2255	5	I04527	I04527	Sequence 1	883	45	2.6	2594	7	CMACON	X82840 C. melo mRNA
C 811	45	2.6	1964	5	I08158	I08158	Sequence 6	884	45	2.6	1395	7	CMGTB	X73314 C. maxima MR
C 812	45	2.6	1964	5	I08168	I08168	Sequence 12	885	45	2.6	1772	7	CREATPS	J04219 C. reinhardt
C 813	45	2.6	876	5	I13140	I13140	Sequence 1	886	45	2.6	1827	7	CRO7332	AJ007332 Catharant
C 814	45	2.6	7859	5	I13185	I13185	Sequence 4	887	45	2.6	774	7	CRRPS18	X83693 C. reinhardt
C 815	45	2.6	2158	5	I17544	I17544	Sequence 1	888	45	2.6	730	7	CRUB13	X60826 C. reinhardt
C 816	45	2.6	2010	5	I17767	I17767	Sequence 4	889	45	2.6	2145	7	CSNADP	X58832 C. sorokinia
C 817	45	2.6	953	5	I19705	I19705	Sequence 1	890	45	2.6	1375	7	EGMDH	X78800 E. gunnii MR
C 818	45	2.6	1493	5	I19901	I19901	Sequence 24	891	45	2.6	3271	7	FSUD1671	U51671 Fusarium SO
C 819	45	2.6	2114	5	I22020	I22020	Sequence 6	892	45	2.6	969	7	GHY9725	AJ009725 Gerbera h
C 820	45	2.6	972	5	I22272	I22272	Sequence 1	893	45	2.6	1754	7	HVADH1	X07774 Barley mRNA
C 821	45	2.6	879	5	I25368	I25368	Sequence 1	894	45	2.6	3292	7	IPBSPP	M64362 Sweet potat
C 822	45	2.6	3200	5	I26205	I26205	Sequence 1	895	45	2.6	1929	7	LEGALX	X99851 L. erecta MR
C 823	45	2.6	630	5	I26207	I26207	Sequence 1	896	45	2.6	957	7	LERNALX	X79338 L. esculentu
C 824	45	2.6	3200	5	I26667	I26667	Sequence 1	897	45	2.6	1032	7	LETHM18GE	X98308 L. esculentu
C 825	45	2.6	857	5	I29004	I29004	Sequence 1	898	45	2.6	923	7	LEUBCI	X82938 L. esculentu
C 826	45	2.6	2384	5	I30334	I30334	Sequence 1	899	45	2.6	2142	7	MCMOD1	X64434 M. crystalli
C 827	45	2.6	2381	5	I30339	I30339	Sequence 6	900	45	2.6	737	7	MTY15368	Y15368 Medicago tr
C 828	45	2.6	1534	5	I33391	I33391	Sequence 6	901	45	2.6	5617	7	MZEEFGLU	M59190 Zea mays fe
C 829	45	2.6	50	5	I36502	I36502	Sequence 1	902	45	2.6	859	7	NPSAPHPRO	X61664 N. sylvestri
C 830	45	2.6	1181	5	I38457	I38457	Sequence 27	903	45	2.6	1657	7	NTACCS	X98492 Nicotiana t
C 831	45	2.6	1493	5	I40996	I40996	Sequence 24	904	45	2.6	1560	7	NTGPRIBET	X98161 N. tabacum m
C 832	45	2.6	1023	5	I41145	I41145	Sequence 16	905	45	2.6	1724	7	NTNKLIN	X77763 N. tabacum N
C 833	45	2.6	1046	5	I44073	I44073	Sequence 4	906	45	2.6	921	7	NTPSAPHP1	X61665 N. tabacum p
C 834	45	2.6	98	5	I49625	I49625	Sequence 42	907	45	2.6	1219	7	NTRIBDR2	X92443 N. tabacum m
C 835	45	2.6	2431	5	I50104	I50104	Sequence 25	908	45	2.6	1138	7	NTIOB	X76188 N. tabacum t
C 836	45	2.6	6671	5	I50973	I50973	Sequence 1	909	45	2.6	1516	7	OSAPHTB1	X78142 O. sativa (A
C 837	45	2.6	1897	5	I52013	I52013	Sequence 1	910	45	2.6	1160	7	OSEMDO	X56063 O. sativa MR
C 838	45	2.6	1181	5	I56932	I56932	Sequence 27	911	45	2.6	1333	7	OSHOX1	X96681 O. sativa MR
C 839	45	2.6	1181	5	I59798	I59798	Sequence 27	912	45	2.6	1685	7	OSTA274	X91608 O. sativa MR
C 840	45	2.6	2223	5	I61429	I61429	Sequence 1	913	45	2.6	858	7	OSU74296	U74296 Oryza sativ
C 841	45	2.6	536	5	I63075	I63075	Sequence 1	914	45	2.6	1148	7	OSY15009	Y15009 Oryza sativ
C 842	45	2.6	2634	5	I6342	I6342	Sequence 1	915	45	2.6	800	7	PCPR2	X58698 P. crispum P
C 843	45	2.6	6671	5	I69486	I69486	Sequence 1	916	45	2.6	2474	7	PEA70HSP	L03399 Pisum sativ
C 844	45	2.6	1181	5	I75125	I75125	Sequence 27	917	45	2.6	1525	7	PEARAT1	D86180 Pisum sativ
C 845	45	2.6	742	5	I77092	I77092	Sequence 12	918	45	2.6	1856	7	PSHDMCA	X73418 P. sativum h
C 846	45	2.6	208	5	I80064	I80064	Sequence 37	919	45	2.6	1656	7	PSHDMCA	X73418 P. sativum h
C 847	45	2.6	1425	5	I81228	I81228	Sequence 1	920	45	2.6	997	7	PSHMGIV	X789568 P. sativum m
C 848	45	2.6	1425	5	I82208	I82208	Sequence 1	921	45	2.6	1596	7	PSVICK	X67429 Pisum sativ
C 849	45	2.6	857	5	I82533	I82533	Sequence 1	922	45	2.6	2036	7	PSY17796	Y17796 Pisum sativ
C 850	45	2.6	2158	5	I82810	I82810	Sequence 1	923	45	2.6	1776	7	RCENOIASE	Z28386 R. communis
C 851	45	2.6	117	5	I89931	I89931	Sequence 3	924	45	2.6	754	7	RICOREB	D10600 Oryza sativ
C 852	45	2.6	69	5	I89934	I89934	Sequence 7	925	45	2.6	764	7	SBNSSSPRT	Y09874 S. bicolor m
C 853	45	2.6	69	5	I89944	I89944	Sequence 22	926	45	2.6	1768	7	SCSACPD	X78935 S. commerson
C 854	45	2.6	2233	5	I92592	I92592	Sequence 1	927	45	2.6	541	7	SHY16247	Y16247 Striga herm
C 855	45	2.6	2384	5	I96058	I96058	Sequence 1	928	45	2.6	1132	7	SLU39059	U39059 Lycopersico
C 856	45	2.6	2381	5	I96063	I96063	Sequence 9	929	45	2.6	2050	7	SOAHRI	X57073 S. oleracea
C 857	45	2.6	1325	5	I96214	I96214	Sequence 51	930	45	2.6	999	7	SOTHRGPB	L22030 Soybean hyd
C 858	45	2.6	723	7	AB015760	AB015760	Nicotiana	931	45	2.6	2070	7	SPLAI	X79705 S. pombe pla
C 859	45	2.6	1364	7	AB025573	AB025573	Nicotiana	932	45	2.6	599	7	TAEPM1	X68289 T. aestivum
C 860	45	2.6	1312	7	AB025715	AB025715	Nicotiana	933	45	2.6	1174	7	TAVDAC2	X82149 T. aestivum
C 861	45	2.6	1978	7	A0ASPSYNM	A0ASPSYNM	A. officinal	934	45	2.6	565	7	TDDEH25	X78320 T. durum Des
C 862	45	2.6	2190	7	ATAAAA	ATAAAA	A. thaliana	935	45	2.6	2564	7	TREXLI	Z69597 T. reesei MR
C 863	45	2.6	1636	7	ATASKDZET	ATASKDZET	A. thaliana	936	45	2.6	2247	7	VFA011303	AJ011303 Vicia fab
C 864	45	2.6	1372	7	ATCOQ3	ATCOQ3	Arabidopsis	937	45	2.6	1146	7	VORNACHIT1	X88800 V. unguicula
C 865	45	2.6	1364	7	ATCSMRNA	ATCSMRNA	A. thaliana	938	45	2.6	972	7	VURNACH14	X88803 V. unguicula
C 866	45	2.6	1649	7	ATH7450	ATH7450	Arabidops	939	45	2.6	7805	7	YSCUC2INDU	L19673 Yeast pYEU1
C 867	45	2.6	1414	7	ATH7587	ATH7587	Arabidops	940	45	2.6	918	7	ZMANB1B	X66077 Z. mays MNB1
C 868	45	2.6	1462	7	ATHRPCA	ATHRPCA	Arabidopsis	941	45	2.6	1048	7	ZMPROIAR	X73379 Z. mays ZMPR
C 869	45	2.6	2923	7	ATHSP91	ATHSP91	A. thaliana	942	45	2.6	1065	7	ZMYPTM2A	X63378 Z. mays ypm
C 870	45	2.6	753	7	ATP29A	ATP29A	A. thaliana	943	45	2.6	1725	8	AF007785	AF007785 Zea mays
C 871	45	2.6	940	7	ATSTZ	ATSTZ	A. thaliana	944	45	2.6	946	8	AF010228	AF010228 Lycopersi
C 872	45	2.6	1050	7	ATTIHPRO	ATTIHPRO	A. thaliana	945	45	2.6	1848	8	AF012862	AF012862 Petroseli
C 873	45	2.6	2357	7	BNAP609	BNAP609	Brassica	946	45	2.6	808	8	AF022741	AF022741 Oryza sat
C 874	45	2.6	2085	7	BOMASPTN	BOMASPTN	X84448 B. oleracea	947	45	2.6	1784	8	AF029856	AF029856 Sorghum b
C 875	45	2.6	1114	7	BRUWIBTCP	BRUWIBTCP	X95640 B. oleracea	948	45	2.6	1319	8	AF031609	AF031609 Oryza sat
C 876	45	2.6	578	7	BRUBIRIPL	BRUBIRIPL	L21898 Brassica ra	949	45	2.6	1380	8	AF038585	AF038585 Zea mays
C 877	45	2.6	567	7	BRUBIRBPA	BRUBIRBPA	Z24738 B. rapa ubiq	950	45	2.6	806	8	AF038949	AF038949 Pinus mon

RESULT	LOCUS	SEQUENCE	FROM PATENT	US	5776457	DNA	PAT	04-DEC-1998
AR015970	AR015970	1737 bp	DNA					
DEFINITION	Sequence 3	from patent	US	5776457				
ACCESSION	AR015970							
NTID	g3972247							
VERSION	AR015970.1	GI:3972247						
KEYWORDS								
SOURCE	Unknown.							
ORGANISM	Unclassified.							
REFERENCE	1 (bases 1 to 1737)							
AUTHORS	Lee,J., and Wood,W.I.							
TITLE	Antibodies to human	PF4A receptor and compositions thereof						
JOURNAL	Patent: US 5776457-A	3 07-JUL-1998.						
FEATURES	location/Qualifiers							
	1..1737							
	/organism="unknown"							

BASE COUNT	454 a	411 c	373 g	499 t
ORIGIN				
Query Match	100.0%;	Score 1737;	DB 5;	Length 1737;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1737;	Conservative	0;	Mismatches	0; Indels 0; Gaps
QY	1	GAATTCATGTGCTGCGGCGCGCGGCGCAAAAGTACGCGGAGGCGCTGAGTCTCCAGTA	60	
Db	1	GAATTCACATGTGCTGCGGCGCGCGCGCAAAAGTACGCGGAGGCGCTGAGTCTCCAGTA	60	
QY	61	GCCACCGCATCTGAGAAACACAGCGGTTCCATGAGGGGGATCACTATATACCTCAGAT	120	
Db	61	GCCACCGCATCTGAGAAACACAGCGGTTCCATGAGGGGGATCACTATATACCTCAGAT	120	
QY	121	AACCTACCCGAGGAAATGGGCTCAGGGGACTATGACTTCATGAGGAACCCGTGTTCCGT	180	
Db	121	AACCTACCCGAGGAAATGGGCTCAGGGGACTATGACTTCATGAGGAACCCGTGTTCCGT	180	
QY	181	GAAAGAAATGCTAAATTCAATAAATCTCTCGCCACCATCTACTCATCATCTTCTTA	240	
Db	181	GAAAGAAATGCTAAATTCAATAAATCTCTCGCCACCATCTACTCATCATCTTCTTA	240	
QY	241	ACTGCGATTGTGGGCAATGAGATTGGTCATCTCTGTCATGAGGTTACGAGAAACTGAGA	300	
Db	241	ACTGCGATTGTGGGCAATGAGATTGGTCATCTCTGTCATGAGGTTACCGAAGAACTGAGA	300	
QY	301	AGCATGACGAGCAATACAGGCTGCACCTGTCACTGAGTGGCGCACTCCTCTTGTCAACG	360	
Db	301	AGCATGACGAGCAATACAGGCTGCACCTGTCACTGAGTGGCGCACTCCTCTTGTCAACG	360	
QY	361	CTTCCCTTCTGGGCGAGTGTGATGCGGTGGCAAACTGATTTGGGAACCTTCATGCAAG	420	
Db	361	CTTCCCTTCTGGGCGAGTGTGATGCGGTGGCAAACTGATTTGGGAACCTTCATGCAAG	420	
QY	421	GCAATTCATGTTCATTCACAGTAACTCTACAGCAGTGTCTCATCTCTGGGCTTCATC	480	
Db	421	GCAATTCATGTTCATTCACAGTAACTCTACAGCAGTGTCTCATCTCTGGGCTTCATC	480	
QY	481	AGCTTGGACCGCTACCTGCGCATCTGTCACAGCCACCAACAGTACAGGGCCAGGAAGCTG	540	
Db	481	AGCTTGGACCGCTACCTGCGCATCTGTCACAGCCACCAACAGTACAGGGCCAGGAAGCTG	540	
QY	541	TTGGCTGAANAAGTGTCTATGTTGGCGCTGTGATCCCTCCCTCTGCTGACTATCCC	600	
Db	541	TTGGCTGAANAAGTGTCTATGTTGGCGCTGTGATCCCTCCCTCTGCTGACTATCCC	600	
QY	601	GACCTCATCTTGGCAACGTCACAGGAGGACAGATGACAGATATATCTGTACCGCTTTCAC	660	
Db	601	GACCTCATCTTGGCAACGTCACAGGAGGACAGATGACAGATATATCTGTACCGCTTTCAC	660	
QY	661	CCCAATGACTTGGGTGTGTTGTTCCATTTACACATCATGATGTTGGCTTATCCG	720	
Db	661	CCCAATGACTTGGGTGTGTTGTTCCATTTACACATCATGATGTTGGCTTATCCG	720	
QY	721	CCTGGTATGTGATCCCTGCTGCTGCTATGTCATTATATCTCCAAAGCTGTACACTCCAG	780	
Db	721	CCTGGTATGTGATCCCTGCTGCTGCTATGTCATTATATCTCCAAAGCTGTACACTCCAG	780	
QY	781	GAGCACCGAAGGCGCAGAGGCTCAAGACACACAGTATCTCATCTGAGCTTTCTTGGC	840	
Db	781	GAGCACCGAAGGCGCAGAGGCTCAAGACACACAGTATCTCATCTGAGCTTTCTTGGC	840	
QY	841	TGTTGGCTGCTTACTACATTTGGAATGAGATGACATGACTCCTTCATCCTCTGGAAATCATC	900	
Db	841	TGTTGGCTGCTTACTACATTTGGAATGAGATGAGATGACTCCTTCATCCTCTGGAAATCATC	900	
QY	901	AAGCAAGGATGTGATTTGGAACACTGTGACAAATGATTTCCATACGAGAGCCCTTA	960	
Db	901	AAGCAAGGATGTGATTTGGAACACTGTGACAAATGATTTCCATACGAGAGCCCTTA	960	
QY	961	GCTTTCTTCACATGTGTCTGAAACCCCATCTCTATGCTTCTTGAGGCCAAATTTTAA	1020	

Db 961 GCTTCTTCACATGTTGTTCTGACCCCAATCCCTATGCTTTCTTGAGCCCAATTTAA 1020
|||
Qy 1021 ACCTCTGCCACGACGACCTACTCTGTGACGAGGGTCCAGCCTCAAGATCTCTCC 1080
|||
Db 1021 ACCTCTGCCACGACGACCTACTCTGTGACGAGGGTCCAGCCTCAAGATCTCTCC 1080
|||
Qy 1081 AAGGAAAGCAGGATGACATTCATCTGTTCCACTGAGTGGAGCTTCAAGTTTAC 1140
|||
Db 1081 AAGGAAAGCAGGATGACATTCATCTGTTCCACTGAGTGGAGCTTCAAGTTTAC 1140
|||
Qy 1141 TCCAGCTAACACAGATGTAAGACCTTTTATATACATTAATTAATTTTAACTT 1200
|||
Db 1141 TCCAGCTAACACAGATGTAAGACCTTTTATATACATTAATTAATTTTAACTT 1200
|||
Qy 1201 ACACATTTTACATATATAAGACCTGACCAATATTTGTACAGTTTATGCTTGTGAT 1260
|||
Db 1201 ACACATTTTACATATATAAGACCTGACCAATATTTGTACAGTTTATGCTTGTGAT 1260
|||
Qy 1261 TTTGCTCTGTTCTTCTTATGTTTGTGAAGTTTAACTTATTAATTAATTTT 1320
|||
Db 1261 TTTGCTCTGTTCTTCTTATGTTTGTGAAGTTTAACTTATTAATTAATTTT 1320
|||
Qy 1321 TTTGCTCTGTTCTTCTTATGTTTGTGAAGTTTAACTTATTAATTTTAACTT 1380
|||
Db 1321 TTTGCTCTGTTCTTCTTATGTTTGTGAAGTTTAACTTATTAATTTTAACTT 1380
|||
Qy 1381 GTCTGCTGAGTGTGAGTGTGAGAAAGGGAAGTGAATCTCCAGAGGTGTATGATCAG 1440
|||
Db 1381 GTCTGCTGAGTGTGAGTGTGAGAAAGGGAAGTGAATCTCCAGAGGTGTATGATCAG 1440
|||
Qy 1441 TAAAGCTAGAAATGATCCCGAGCTGTTATGATAGATTAATCTCTCCATCCCGTGAA 1500
|||
Db 1441 TAAAGCTAGAAATGATCCCGAGCTGTTATGATAGATTAATCTCTCCATCCCGTGAA 1500
|||
Qy 1501 GTTTTCTCTGTTCTTAAAGACCTGATTTTGCTGTAGAAATGATGACCTTAATCAAGCCC 1560
|||
Db 1501 GTTTTCTCTGTTCTTAAAGACCTGATTTTGCTGTAGAAATGATGACCTTAATCAAGCCC 1560
|||
Qy 1561 AAGTGTATGAAATGCTGTTTTCAGTTTTCAGAGTGGTGTATGAGACCTAC 1620
|||
Db 1561 AAGTGTATGAAATGCTGTTTTCAGTTTTCAGAGTGGTGTATGAGACCTAC 1620
|||
Qy 1621 AGTGTACAGTCTTGTATTAAGTGTATTAATAAAGTAACTTTAACTTAATAAAAA 1680
|||
Db 1621 AGTGTACAGTCTTGTATTAAGTGTATTAATAAAGTAACTTTAACTTAATAAAAA 1680
|||
Qy 1681 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCCGACACACTGGAATTC 1737
|||
Db 1681 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCCGACACACTGGAATTC 1737
|||

Query Match 100.0%; Score 1737; DB 5; Length 1737;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1737; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GAATTCAGTGTGTTGGGGGGGGGCAAAAGTACACCCGAGGCGCTGAGTGTCCAGTA 60
|||
Db 1 GAATTCAGTGTGTTGGGGGGGGGCAAAAGTACACCCGAGGCGCTGAGTGTCCAGTA 60
|||
Qy 61 GCCACCGCATCTGGAAGAACACGCGTTACCATGAGGAGGATCACTATATACACTCAGAT 120
|||
Db 61 GCCACCGCATCTGGAAGAACACGCGTTACCATGAGGAGGATCACTATATACACTCAGAT 120
|||
Qy 121 AACTACACCGAGAAATGGGCTCAGGGAGTATGATCACTCAAGAAAGAACCTGTTCCGT 180
|||
Db 121 AACTACACCGAGAAATGGGCTCAGGGAGTATGATCACTCAAGAAAGAACCTGTTCCGT 180
|||
Qy 181 GAAGAAATGCTAATTTCAATAAATCTCTCGCCACATCTACTCATCATCTTCTTA 240
|||
Db 181 GAAGAAATGCTAATTTCAATAAATCTCTCGCCACATCTACTCATCATCTTCTTA 240
|||
Qy 241 ACTGCGATGTGGGCAATGATTTGCTCATCTGCTCATGCGTTACCAAGAAACTGAGA 300
|||
Db 241 ACTGCGATGTGGGCAATGATTTGCTCATCTGCTCATGCGTTACCAAGAAACTGAGA 300
|||
Qy 301 AGCATGAGGACAGTACAGGCTGACCTGTCAATGAGCCGCTCTTGTGATCAG 360
|||
Db 301 AGCATGAGGACAGTACAGGCTGACCTGTCAATGAGCCGCTCTTGTGATCAG 360
|||
Qy 361 CTTCCTCTTGGGCAATGATTTGCTGCAAACTGATTTGGAACCTTCTATGCAAG 420
|||
Db 361 CTTCCTCTTGGGCAATGATTTGCTGCAAACTGATTTGGAACCTTCTATGCAAG 420
|||
Qy 421 GCAATCATGATCTACACAGTAACTTACACAGTGTCTCATCTGCGCTTATC 480
|||
Db 421 GCAATCATGATCTACACAGTAACTTACACAGTGTCTCATCTGCGCTTATC 480
|||
Qy 481 AGTGTGACGCTACCTGCGCATGCTCCAGCCCAACAGTGTGAGAGCCAAAGAACTG 540
|||
Db 481 AGTGTGACGCTACCTGCGCATGCTCCAGCCCAACAGTGTGAGAGCCAAAGAACTG 540
|||
Qy 541 TTGGCTGAAAGGCTGCTATGTTGGGCTGATCCCTGCTGCTGATATTC 600
|||
Db 541 TTGGCTGAAAGGCTGCTATGTTGGGCTGATCCCTGCTGCTGATATTC 600
|||
Qy 601 GACTTCATCTTGGCAACGTCAGTGAAGGAGATGACAGATATATGTGTACCCGCTTAC 660
|||
Db 601 GACTTCATCTTGGCAACGTCAGTGAAGGAGATGACAGATATATGTGTACCCGCTTAC 660
|||
Qy 661 CCCAATGACTTGGGAGTGTGTTGTTCCATTTCAGACATCAATGAGTGGCTTATCTG 720
|||
Db 661 CCCAATGACTTGGGAGTGTGTTGTTCCATTTCAGACATCAATGAGTGGCTTATCTG 720
|||
Qy 721 CTTGGTATTTGTCATCTGCTCTGCTATTCGATTATCATCTCAAGCTGTACACTCCAG 780
|||
Db 721 CTTGGTATTTGTCATCTGCTCTGCTATTCGATTATCATCTCAAGCTGTACACTCCAG 780
|||
Qy 781 GGGCACCAGAGGCAAGGCGCTCAAGACACAGTCACTCTCATCTCGGCTTCTTCCG 840
|||
Db 781 GGGCACCAGAGGCAAGGCGCTCAAGACACAGTCACTCTCATCTCGGCTTCTTCCG 840
|||
Qy 841 TGTGGCTGCTTACTCATTTGGATGACGATGACTCTTCACTCTCGAAGATCATC 900
|||
Db 841 TGTGGCTGCTTACTCATTTGGATGACGATGACTCTTCACTCTCGAAGATCATC 900
|||
Qy 901 AAGCAAGGCTGTGATTTGGAACACCTGTGACAGTGTGATTTTCATCAACGAGGCTTA 960
|||
Db 901 AAGCAAGGCTGTGATTTGGAACACCTGTGACAGTGTGATTTTCATCAACGAGGCTTA 960
|||
Qy 961 GCTTCTTCCAGTGTGTGTGAAACCCATCTATGCTTCTTGGAGCAATTTAA 1020
|||
Db 961 GCTTCTTCCAGTGTGTGTGAAACCCATCTATGCTTCTTGGAGCAATTTAA 1020
|||

	Matches 1/1/	Conservative	0/	Matches	0/
QY	1	GAATTCAGAGTCTGGCGCGCGGCGCAAAATGACGCGCGAGGCTGAGTGGCTCCAGTA	60		
Db	1	GAATTCAGAGTCTGGCGCGCGCGCGCAAAATGACGCGCGAGGCGCTGAGTGGCTCCAGTA	60		
QY	61	GCCACCGCATCTGGAAACCAACCGGTTACCATGAGGGGATCAGATATACACTTCAGAT	120		
Db	61	GCCACCGCATCTGGAAACCAACCGGTTACCATGAGGGGATCAGATATACACTTCAGAT	120		
QY	121	AACATACCCGAGAAATGGGCTCAGGGGACTATGACTCATGAAGAAACCTGTTCCGT	180		
Db	121	AACATACCCGAGAAATGGGCTCAGGGGACTATGACTCATGAAGAAACCTGTTCCGT	180		
QY	181	GAAAGAAATGCTAATTTCAATAAATCTTCGCGCCACACATCATCATCATCTCTTA	240		
Db	181	GAAAGAAATGCTAATTTCAATAAATCTTCGCGCCACACATCATCATCATCTCTTA	240		
QY	241	ACTGGCATGAGGCAATGATTTGTCATCTGTCATGATGGTTTACCAGAAAGAACTGAGA	300		
Db	241	ACTGGCATGAGGCAATGATTTGTCATCTGTCATGATGGTTTACCAGAAAGAACTGAGA	300		
QY	301	AGCATGACGACAAATACAGGCTGCACCTGTAGTGGCGGACCTCTCTTTGATATCAG	360		
Db	301	AGCATGACGACAAATACAGGCTGCACCTGTAGTGGCGGACCTCTCTTTGATATCAG	360		
QY	361	CTTCCCTCTGGGCAATGATGCGCGTGGCAAACTGCTACTTTGGGAATCTCTATGCAAG	420		
Db	361	CTTCCCTCTGGGCAATGATGCGCGTGGCAAACTGCTACTTTGGGAATCTCTATGCAAG	420		
QY	421	GCATGTCATGTCATCTACACAGTCACTTAACGACAGTCTCTAATCTGGCCTTTCATC	480		
Db	421	GCATGTCATGTCATCTACACAGTCACTTAACGACAGTCTCTAATCTGGCCTTTCATC	480		
QY	481	ACTGCGACCGTACCTGGCCATGTCACAGCCACCAACAGTCAGAGGCCAAGGAAGCTG	540		
Db	481	ACTGCGACCGTACCTGGCCATGTCACAGCCACCAACAGTCAGAGGCCAAGGAAGCTG	540		
QY	541	TTGGGCTGAAAAGATGCTATGTTGGCGTGTGGATCCCTGCTCTGCTGACTATGCC	600		
Db	541	TTGGGCTGAAAAGATGCTATGTTGGCGTGTGGATCCCTGCTCTGCTGACTATGCC	600		
QY	601	GACTTCATCTTGGCCAAAGTCAGTGAAGGACAGATACAGATATATCTGTGACCGCTTCTAC	660		
Db	601	GACTTCATCTTGGCCAAAGTCAGTGAAGGACAGATACAGATATATCTGTGACCGCTTCTAC	660		
QY	661	CCCATGACTTGGGTTGTTGTTCCAGTTTCAGCATATCATGTTGGCTTATCTTG	720		
Db	661	CCCATGACTTGGGTTGTTGTTCCAGTTTCAGCATATCATGTTGGCTTATCTTG	720		
QY	721	CCGATGATGTCATCTGTCGTCGATGATGATATTCATCTCCAAAGCTGTACACTCCAAAG	780		
Db	721	CCGATGATGTCATCTGTCGTCGATGATGATATTCATCTCCAAAGCTGTACACTCCAAAG	780		
QY	781	GCGCACAGAAAGCGAAAGGCCCTCAAGACCAAGTCATCTCATCTGCTTCTTCCGC	840		
Db	781	GCGCACAGAAAGCGAAAGGCCCTCAAGACCAAGTCATCTCATCTGCTTCTTCCGC	840		
QY	841	TGTTGGCTGCTTACTACATTTGGGATCGAATGCAATCCTTCATCTCCCTGGAATATATC	900		
Db	841	TGTTGGCTGCTTACTACATTTGGGATCGAATGCAATCCTTCATCTCCCTGGAATATATC	900		
QY	901	AAGCAAGGGTGTAGTTTGAAGAACTGTGACAAAGTGGATTTCCATACCGGAGGCTTA	960		
Db	901	AAGCAAGGGTGTAGTTTGAAGAACTGTGACAAAGTGGATTTCCATACCGGAGGCTTA	960		
QY	961	GCTTCTTCCATGTTGTCTGAACCCCATCTCTATGCTTCTCTTGAGGCCAAATTTTAAA	1020		
Db	961	GCTTCTTCCATGTTGTCTGAACCCCATCTCTATGCTTCTCTTGAGGCCAAATTTTAAA	1020		
QY	1021	ACCTTCGCGAGGACGACTCACTGTGAGCAAGAGGTCAGGCTCAAGATCCTCTCC	1080		
Db	1021	ACCTTCGCGAGGACGACTCACTGTGAGCAAGAGGTCAGGCTCAAGATCCTCTCC	1080		


```

Db 664 CCAAGTTTCAGACATCATGGTTGGCCTTATCTGCTGCTGATATGTCATCTGCTGCTA 723
QY 747 TTGCATTATCATCTCCAGCTGTACACCTCCAGAGGCCACAGAGCGAAGGCCCTCAA 806
Db 724 TTGCATTATCATCTCCAGCTGTACACCTCCAGAGGCCACAGAGCGAAGGCCCTCAA 783
QY 807 GACCACAGTCATCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 866
Db 784 GACCACAGTCATCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 843
QY 867 GACCACAGTCATCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 926
Db 844 CACATCTGACTCTCTATCTCTCTGGAATTCATCAAGAGAGGAGTGTAGATGAGAAC 903
QY 927 TGTGCAAGTGTGATTTTCATCAACGAGGCCCTAGCTTTCTTCTCACTGTTGTGTAACC 986
Db 904 TGTGCAAGTGTGATTTTCATCAACGAGGCCCTAGCTTTCTTCTCACTGTTGTGTAACC 963
QY 987 CATCTCTATGCTTCTCTGAGAGCCAAATTTAAACCTCTGCCAGACGACTACCTC 1046
Db 964 CATCTCTATGCTTCTCTGAGAGCCAAATTTAAACCTCTGCCAGACGACTACCTC 1023
QY 1047 TGTGACACAGAGGCTCAGGCTCAAGATCCTCTCCAAAGGAAAGGAGGAGATTCATC 1106
Db 1024 TGTGACACAGAGGCTCAGGCTCAAGATCCTCTCCAAAGGAAAGGAGGAGATTCATC 1083
QY 1107 TGTTCACAGTGTGAGTCTCAAGTCTCAAGTCTTCACTCCAGCTAACAGAGATGTAAG 1166
Db 1084 TGTTCACAGTGTGAGTCTCAAGTCTTCAAGTCTTCACTCCAGCTAACAGAGATGTAAG 1143
QY 1167 TTTTATATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1226
Db 1144 TTTTATATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1203
QY 1227 ACCAATATGTAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1286
Db 1204 ACCAATATGTAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1263
QY 1287 GTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1346
Db 1264 GTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1323
QY 1347 GCAGAGCTGTGAGCCAAAGTCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1406
Db 1324 GCAGAGCTGTGAGCCAAAGTCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1383
QY 1407 GAACGATGATTCAGAGGCTGTAGTAATCAGCTAAGGTAAGTAAATGCCAGCTGT 1466
Db 1384 GAACGATGATTCAGAGGCTGTAGTAATCAGCTAAGGTAAGTAAATGCCAGCTGT 1443
QY 1467 TTATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1526
Db 1444 TTATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1503
QY 1527 TTGCTGTAGAGTGTGACTTATTAACCAAGGCCAAAGTGTATTAAGATGCTGTTT 1586
Db 1504 TTGCTGTAGAGTGTGACTTATTAACCAAGGCCAAAGTGTATTAAGATGCTGTTT 1563
QY 1587 CAGTTTTCAGAGTGTGACTTATTAACCAAGGCCAAAGTGTATTAAGATGCTGTT 1646
Db 1564 CAGTTTTCAGAGTGTGACTTATTAACCAAGGCCAAAGTGTATTAAGATGCTGTT 1623
QY 1647 AATTAAGTACATGTTAACTT 1668
Db 1624 AATTAAGTACATGTTAACTT 1645

```

```

VERSION D10924.1 GI:219868
KEYWORDS GTP-binding protein; plasma membrane protein; protein coupled.
SOURCE Homo sapiens monocyte, CDNA to mRNA.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1664)
AUTHORS Nomura,H., Nielsen,B.W. and Matsushima,K.
TITLE Molecular cloning of cDNAs encoding a LD78 receptor and putative leukocyte chemotactic peptide receptors
JOURNAL Int. Immunol. 5 (10), 1239-1249 (1993)
MEDLINE 94092629
COMMENT Submitted (13-Apr-1992) to DDBJ by: Hideki Nomura
Dept. of Pharmacol.
Cancer Res. Inst., Kanazawa Univ.
13-1 Takaramachi
Kanazawa, Ishikawa 920
Japan
Phone: 0762-62-8151 x5875
Fax: 0762-60-7704.
FEATURES
source
1..1664
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="human monocyte"
/cell_type="monocyte"
/gene="HM89"
77..1135
/gene="HM89"
77..1135
/codon_start=1
/product="HM89"
/protein_id="BAA01722.1"
/db_xref="PID:d1002197"
/db_xref="PID:g219869"
/db_xref="PID:g219869"
/db_xref="GI:219869"
/translacion="MEGISITVSNYEDMSGDYDSKKEPCFRENANENKILPIL
YSIFLNGIVNGVILVMGYOKLRSTDRYRLSVADLLFYITLPEFVADVAVAM
YEGNFKCAVAVIVLVYSLVLLAFSLDRYLAIVATNSQRPRLAEKVVYGV
MIPLLILIPFIFANVSEADRYICDFEYNDLMVYVFOHIMVGLILGIVLSC
YCLITRSKSHRGOKRKALKTYYLILAFACMLPYIGISISFILLETIKGCEP
ENTYHKMISTEDALAFRHCNPLILYATLGLAKRTSAOHALTVSRGSLKILSKGR
GGHSSVSSESSSFHSS"
BASE COUNT 414 a 398 c 359 g 493 t
ORIGIN
Query Match 92.2%; Score 1601; DB 9; Length 1664;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1651; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 27 GCAAGTGAAGCGGAGGCGCTGAGTGCCTCAGTACGACCCGATCTGGAGACGAGCGT 86
Db 13 GCAAGTGAAGCGGAGGCGCTGAGTGCCTCAGTACGACCCGATCTGGAGACGAGCGT 72
QY 87 TACCATGAGAGGAGATCAGTATATACACTTCAGATTAACACCGAGAAATGGGCTCAG 146
Db 73 TACCATGAGAGGAGATCAGTATATACACTTCAGATTAACACCGAGAAATGGGCTCAG 132
QY 147 GGAATATGACTCCATGAGAGAACCTGTTTCGCTAAGAGAAATGCTAATTTAATTAAT 206
Db 133 GGAATATGACTCCATGAGAGAACCTGTTTCGCTAAGAGAAATGCTAATTTAATTAAT 192
QY 207 CTTCCTGGCCACACATCTACTCATCATCTCTTAACTGGCATGTGGCAATGGATGGT 266
Db 193 CTTCCTGGCCACACATCTACTCATCATCTCTTAACTGGCATGTGGCAATGGATGGT 252
QY 267 CATCTGTGATGAGGTATACAGAGAAACTGAGAGCATGACGAGCAAGTACAGGCTGCA 326
Db 253 CATCTGTGATGAGGTATACAGAGAACTGAGAGCATGACGAGCAAGTACAGGCTGCA 312
QY 327 CCTGTGATGAGGCGGAGACCTCTTGTGTATCAACGCTTCCCTTCTGGGCAAGTATGCCGT 386

```

```

RESULT 5
HUMH89 HUMH89 1664 bp mRNA PRI 03-FEB-1999
DEFINITION Human mRNA for HM89.
ACCESSION D10924
FEATURES
    ORIGIN
        9219868

```



```

Db 313 CCTGTGAGTGGCGACCTCCCTCTTGTATCACCCTTCCCTTGGGAGTTGATGCCGT 372
QY 387 GCGAATATGTTACTTTGGGAATCTTCTATGCAAGGAGTCCATCTCATACAGTCAA 446
Db 373 GCGAATATGTTACTTTGGGAATCTTCTATGCAAGGAGTCCATCTCATACAGTCAA 432
QY 447 CCTGTACAGAGTGTCTCTCATCTCTGCGCTTCATAGTGTGAGCGGTACCTGGGCATGT 506
Db 433 CCTGTACAGAGTGTCTCTCATCTCTGCGCTTCATAGTGTGAGCGGTACCTGGGCATGT 492
QY 507 CCAAGCCACCAAGAGTGAAGGCCAAGAAAGCTTGTGGCTGAAAAGGTGTCTATGTGG 566
Db 493 CCAAGCCACCAAGAGTGAAGGCCAAGAAAGCTTGTGGCTGAAAAGGTGTCTATGTGG 552
QY 567 CGTGTGATCCCTGCGCTCTCTGCTGATATTCGGACTTCATCTTGGCAAGCTCATGA 626
Db 553 CGTGTGATCCCTGCGCTCTCTGCTGATATTCGGACTTCATCTTGGCAAGCTCATGA 612
QY 627 GCGAGATGACAGATATATCTGTGACCGCTTCTACCCCAATGACTTGGGTGTGTGT 686
Db 613 GCGAGATGACAGATATATCTGTGACCGCTTCTACCCCAATGACTTGGGTGTGTGT 672
QY 687 CCAAGTTTACACATCATGTGTGGCTTATCTGCTGCTGCTGATTTCTATCTCTCTCTA 746
Db 673 CCAAGTTTACACATCATGTGTGGCTTATCTGCTGCTGCTGATTTCTATCTCTCTCTA 732
QY 747 TTGCATATATCTCTCCAGGTGTCTACCTGCAAGGCGCACAGAGGCGCAAGGCCCTCAA 806
Db 733 TTGCATATATCTCTCCAGGTGTCTACCTGCAAGGCGCACAGAGGCGCAAGGCCCTCAA 792
QY 807 GACCACAGTATCTCTCACTCTGCTGCTTCTGCGCTGTGGCTGCTTACTACTATTTGGAT 866
Db 793 GACCACAGTATCTCTCACTCTGCTGCTTCTGCGCTGTGGCTGCTTACTACTATTTGGAT 852
QY 867 CACATCTGATCTCTTCTATCTCTCTGCAAAATCATCAACAAAGGCTGTAGTTTGAACAC 926
Db 853 CACATCTGATCTCTTCTATCTCTCTGCAAAATCATCAACAAAGGCTGTAGTTTGAACAC 912
QY 927 TGTGCAAGTGTGATTTCCATCTACCGAGGCGCTAGCTTTCTTCCACTGTGTGTGAACCC 986
Db 913 TGTGCAAGTGTGATTTCCATCTACCGAGGCGCTAGCTTTCTTCCACTGTGTGTGAACCC 972
QY 987 CATCTCTATGCTTTCTCTTGGAGCCAAATTTAAACCTCTGCCAGACGACATCACCTC 1046
Db 973 CATCTCTATGCTTTCTCTTGGAGCCAAATTTAAACCTCTGCCAGACGACATCACCTC 1032
QY 1047 TGTGACAGAGGCTCCAGCTCAAGATCTCTCCAAAGGAAAGGAGGTGACATTTGATC 1106
Db 1033 TGTGACAGAGGCTCCAGCTCAAGATCTCTCCAAAGGAAAGGAGGTGACATTTGATC 1092
QY 1107 TGTTCACAGTGTGAGTCTTCAAGTTTCTCACTCCAGCTACACAGATGTAAAGACT 1166
Db 1093 TGTTCACAGTGTGAGTCTTCAAGTTTCTCACTCCAGCTACACAGATGTAAAGACT 1152
QY 1167 TTTTATTAACGATTAATAAATTTTTTTTAAGTTACACATTTTTCAGATATATAAGACTG 1226
Db 1153 TTTTATTAACGATTAATAAATTTTTTTTAAGTTACACATTTTTCAGATATATAAGACTG 1212
QY 1227 ACCAATATGTACAGTTTATGCTGTGGAATTTTGTCTGTGTTCTTGTAGTTTT 1286
Db 1213 ACCAATATGTACAGTTTATGCTGTGGAATTTTGTCTGTGTTCTTGTAGTTTT 1272
QY 1287 GTGAGTTTAAATGACTTATTTAATAAATTTTTTTTGTTCATATGATGTGTCTAG 1346
Db 1273 GTGAGTTTAAATGACTTATTTAATAAATTTTTTTTGTTCATATGATGTGTCTAG 1332
QY 1347 GCGAGACCTGTGGCAAGTTCTAGTGTCTGTATGCTCTGTGTAGACCTGTGAAAAGG 1406
Db 1333 GCGAGACCTGTGGCAAGTTCTAGTGTCTGTATGCTCTGTGTAGACCTGTGAAAAGG 1392
QY 1407 GAACCTAACATTCAGAGCTGTAGTGAATCAGGTAAGCTAGAAAAGATCCCAAGCTGT 1466

```

```

Db 1393 GAACCTAACATTCAGAGCTGTAGTGAATCAGGTAAGCTAGAAAAGATGATCCCAAGCTGT 1452
QY 1467 TTATGATAGATATATCTCTCTCATTTCCCGTGAAGAGTTTTCCTGTTTAAAGACTATTT 1526
Db 1453 TTATGATAGATATATCTCTCTCATTTCCCGTGAAGAGTTTTCCTGTTTAAAGACTATTT 1512
QY 1527 TTGCTGTAGAAGATGGCACTTATTAACCAAGCCCAAGGTGATAGAAATGCTGTTT 1586
Db 1513 TTGCTGTAGAAGATGGCACTTATTAACCAAGCCCAAGGTGATAGAAATGCTGTTT 1572
QY 1587 CAGTTTCAAGAGTGGTGTGATTTACAGCACCCTACAGTACAGTCTGTGTAATTAAGTTGT 1646
Db 1573 CAGTTTCAAGAGTGGTGTGATTTACAGCACCCTACAGTACAGTCTGTGTAATTAAGTTGT 1632
QY 1647 AATAAAGTACATGTTAACTTAAAAA 1678
Db 1633 AATAAAGTACATGTTAACTTAAAAA 1664

```

```

RESULT 6
LOCUS 628514 1664 bp DNA STS 01-JUL-1996
DEFINITION human STS SHGC-31461, sequence tagged site.
ACCESSION 628514
NID 91408329
VERSION 628514.1 GI:1408329
KEYWORDS STS: STS sequence; primer; sequence tagged site.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiinae; Homo.
1 (bases 1 to 1664)
REFERENCE
AUTHORS Myers,R.M.
JOURNAL Unpublished (1996)
COMMENT

```

Contact: Richard M. Myers
Stanford Human Genome Center (SHGC)
Stanford University School of Medicine
Department of Genetics, M-344, Stanford, CA 94305, USA
Tel: 4157259687
Fax: 4157259689
Email: myers@shgc.stanford.edu

Primer A: TTGTACAGTTTATTTGCTTGTG
Primer B: CGTGATTAACACAGCTCTGG
STS size: 207
PCR Profile:
Initial incubation: 94 degrees C for 90 seconds

Denaturation: 94 degrees C for 15 seconds
Annealing: 62 degrees C for 23 seconds
Polymerization: 72 degrees C for 30 seconds
PCR Cycles: 30
Thermal Cycler: Perkin Elmer 9600

Protocol:
Template: 25 ng
Primer: each 1 uM
dNTPs: each 200 uM
Taq Polymerase: 0.05 units/uL
Total Vol: 10 uL

Buffer:
MgCl2: 2.5 mM
KCl: 50 mM
Tris-HCl: 20 mM
pH: 8.3

Prepared with primer pairs provided by Sandoz, derived from D10924
-- Washington University/Merck EST sequence.
Location/Qualifiers
1. 1664
/organism="Homo sapiens"
/db_xref="taxon:9606"

FEATURES
source

STS /map="2"
 1220.1426
 primer_bind 1220.1244
 complement(1405.1426)
 BASE COUNT 414 a 398 c 359 g 493 t
 ORIGIN

Query Match 92.2% Score 1601; DB 14; Length 1664;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1651; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 27 GCAAGTACGCGGAGGCTGAGTGTCCAGTAGCCACCGCATCTGGAGAACCGGCT 86
 Db 13 GCAAGTACGCGGAGGCTGAGTGTCCAGTAGCCACCGCATCTGGAGAACCGGCT 72
 Oy 87 TACCATGAGGAGATCATATATACACTCAGATTAACCCAGAGAAATGGCTAGG 146
 Db 73 TACCATGAGGAGATCATATATACACTCAGATTAACCCAGAGAAATGGCTAGG 132
 Oy 147 GAGCTATGATCCATGAGAGAACCTGTTCCGTAAGAAATGCTAATTTCAATAAAT 206
 Db 133 GAGCTATGATCCATGAGAGAACCTGTTCCGTAAGAAATGCTAATTTCAATAAAT 192
 Oy 207 CTTCCTGCCACCATCTACTCATCATCTTCTTAATGCGATTGTGGCAATGGATTGT 266
 Db 193 CTTCCTGCCACCATCTACTCATCATCTTCTTAATGCGATTGTGGCAATGGATTGT 252
 Oy 267 CATCTGCTGATGGGTATACCAAGAAAGATGAGAGATGACGACACAGTACAGCTGCA 326
 Db 253 CATCTGCTGATGGGTATACCAAGAAAGATGAGAGATGACGACACAGTACAGCTGCA 312
 Oy 327 CTCTGCTGATGGGACCTCTCTTTCATCAGCTCCCTTCGGGAGTTGATGGCGT 386
 Db 313 CTCTGCTGATGGGACCTCTCTTTCATCAGCTCCCTTCGGGAGTTGATGGCGT 372
 Oy 387 GGCAGATGCTTGTGGAACTTCTTAAGCAAGGAGTCCATGTCATCTACACAGTCAA 446
 Db 373 GGCAGATGCTTGTGGAACTTCTTAAGCAAGGAGTCCATGTCATCTACACAGTCAA 432
 Oy 447 CCTCTACAGAGTGTCTCTATCCTGGGCTTCATCAGTCTGGACCGCTACTGGCATCGT 506
 Db 433 CCTCTACAGAGTGTCTCTATCCTGGGCTTCATCAGTCTGGACCGCTACTGGCATCGT 492
 Oy 507 CCACGCCACCAAGTACAGAGGCCCAAGAGCTGTGGTGAAGAGTGTCTATGTTGG 566
 Db 493 CCACGCCACCAAGTACAGAGGCCCAAGAGCTGTGGTGAAGAGTGTCTATGTTGG 552
 Oy 567 CGTCTGGATCCCTGCTCTCTGCTGACTATTCGGACTTATCTTTGCCAAGCTCACTGA 626
 Db 553 CGTCTGGATCCCTGCTCTCTGCTGACTATTCGGACTTATCTTTGCCAAGCTCACTGA 612
 Oy 627 GCGAGATGAGAGATATATCTGTGACCGCTTCTACCCCAATGACTTGTGGTGTGTT 686
 Db 613 GCGAGATGAGAGATATATCTGTGACCGCTTCTACCCCAATGACTTGTGGTGTGTT 672
 Oy 687 CCAATTACAGACATCATGTGTGGCTTATCTGCTGCTGATGATTTGATCTGCTGCTGA 746
 Db 673 CCAATTACAGACATCATGTGTGGCTTATCTGCTGCTGATGATTTGATCTGCTGCTGA 732
 Oy 747 TTGCATTATCATCTCAAGCTGTCTACACTCCAAAGGGCCACCAAGGGAGGCCCTCAA 806
 Db 733 TTGCATTATCATCTCAAGCTGTCTACACTCCAAAGGGCCACCAAGGGAGGCCCTCAA 792
 Oy 807 GACACAGTATCTCTATCTGTGCTTCTGTGCTGTGGCTGCTTACTACTATGGGAT 866
 Db 793 GACACAGTATCTCTATCTGTGCTTCTGTGCTGTGGCTGCTTACTACTATGGGAT 852
 Oy 867 CAGCATGACATCTCTCTCTCTCTGTAATCATCAAGCAAGAGGTGTAGTTGAGACAC 926
 Db 853 CAGCATGACATCTCTCTCTCTCTGTAATCATCAAGCAAGAGGTGTAGTTGAGACAC 912
 Oy 927 TGTGCAAGAGTGAATTTTCATCAACGAGGCCCTAGCTTTCTTCCACTGTGTGTGAACCC 986

Db 913 TGTGCAAGAGTGAATTTTCATCAACGAGGCCCTAGCTTTCTTCCACTGTGTGTGAACCC 972
 Oy 987 CATCTCTATGCTTTCTCTGGAGCCAAATTTAAACCTCTGCCACGACCTCACTC 1046
 Db 973 CATCTCTATGCTTTCTCTGGAGCCAAATTTAAACCTCTGCCACGACCTCACTC 1032
 Oy 1047 TGTGAGAGAGGTTCCAGCTCAAGATCTCTCCAAAGGAAAGCGAGTGGACATTCATC 1106
 Db 1033 TGTGAGAGAGGTTCCAGCTCAAGATCTCTCCAAAGGAAAGCGAGTGGACATTCATC 1092
 Oy 1107 TGTTCCTACGATCTGAGTCTTCAAGTTTTCCTCAGCTAACACAGATGTAAAGACT 1166
 Db 1093 TGTTCCTACGATCTGAGTCTTCAAGTTTTCCTCAGCTAACACAGATGTAAAGACT 1152
 Oy 1167 TTTTATATGATTAATACTTTTAAAGTACACATTTTCAGATTAATAAGACTG 1226
 Db 1153 TTTTATATGATTAATACTTTTAAAGTACACATTTTTCAGATTAATAAGACTG 1212
 Oy 1227 ACCAATATGATTAAGTATTTATGCTTGTGGATTTTGTCTTGTCTTATGTTTTT 1286
 Db 1213 ACCAATATGATTAAGTATTTATGCTTGTGGATTTTGTCTTGTCTTATGTTTTT 1272
 Oy 1287 GTGAGTTTATGATTAATTAATAATTTTGTGTTTCAATGATGTGTCTAG 1346
 Db 1273 GTGAGTTTATGATTAATTAATAATTTTGTGTTTCAATGATGTGTCTAG 1332
 Oy 1347 GCAAGACCTGTGGCAAGTCTTAGTGTGATGCTGTGTGTAGGACGTGAAGAAG 1406
 Db 1333 GCAAGACCTGTGGCAAGTCTTAGTGTGATGCTGTGTGTAGGACGTGAAGAAG 1392
 Oy 1407 GAACTGACATTCAGAGGCTGTAGTGAATCAGCTAAAGCTAAGATGATCCACACTGT 1466
 Db 1393 GAACTGACATTCAGAGGCTGTAGTGAATCAGCTAAAGCTAAGATGATCCACACTGT 1452
 Oy 1467 TTAATGATTAATATCTCTCATTCCCGTGAAGCTTTTCTGTTCTTAAGCGTGAAT 1526
 Db 1453 TTAATGATTAATATCTCTCATTCCCGTGAAGCTTTTCTGTTCTTAAGCGTGAAT 1512
 Oy 1527 TTGCTGTAGAGATGGGACCTTATTAACCAAGGCCAAGGTTATGAATGCTGGTTTTT 1586
 Db 1513 TTGCTGTAGAGATGGGACCTTATTAACCAAGGCCAAGGTTATGAATGCTGGTTTTT 1572
 Oy 1587 CAGTTTTCAGAGATGGGATGATTTTTCAGCACTACAGTGAACGCTTGTGAATGTTGT 1646
 Db 1573 CAGTTTTCAGAGATGGGATGATTTTTCAGCACTACAGTGAACGCTTGTGAATGTTGT 1632
 Oy 1647 AATTAAGTACATGTTAACTTAAAAA 1678
 Db 1633 AATTAAGTACATGTTAACTTAAAAA 1664

RESULT 7
 AF005058
 LOCUS 5161 bp DNA PRI 01-JAN-1998
 DEFINITION Homo sapiens chemokine receptor (CXCR-4) gene, complete cds.
 ACCESSION AF005058
 NID 92735718
 VERSION AF005058.1 GI:2735718
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 5161)
 AUTHORS Wegner,S.A., Ehrenberg,P.K., Chang,G., Dayhoff,D.E. and Michael,N.L.
 TITLE Genomic organization and characterization of the promoter for the HIV-1 entry co-receptor CXCR-4
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 5161)
 AUTHORS Wegner,S.A., Ehrenberg,P.K., Chang,G., Dayhoff,D.E. and Michael,N.L.

TITLE Direct Submission
JOURNAL Submitted (21-MAY-1997) Division of Retrovirology, Walter Reed Army
Institute of Research, 13 Taft Court, Suite 200, Rockville, MD
20850, USA

FEATURES Location/Qualifiers
source 1. .5161
/organism="Homo sapiens"
/db_xref="taxon:9606"
2. .5161
/gene="CXCR-4"
TATA_signal 519. .923
/gene="CXCR-4"
exon 950. .1052
/gene="CXCR-4"
5'UTR 950. .1037
/gene="CXCR-4"
join(950. .1052,3185. .4760)
mRNA /gene="CXCR-4"
/product="chemokine receptor"
join(1038. .1052,3185. .4228)
CDS /gene="CXCR-4"
/note="seven transmembrane G-coupled receptor for the CXK
chemokine SDF-1; entry co-receptor for HIV-1 strains with
the syncytium-inducing phenotype"
/codon_start=1
/product="chemokine receptor"
/protein_id="A893982.1"
/db_xref="PID:q2735719"
/db_xref="GI:2735719"
/translation="MEGISLYTSDNYTEKSGSDYSMEKPEFRENNANFNKLFPTI
YSIIFFGVNGLVILVMGYOKRLSMDYKRLHSVADLFLVITLPEVAADVAMV
YGFNFKCAVHIVLYLVNLSVLIILAFISLDRIYLVHATNSGRPKLLAEKVYVYV
WIPALLITIPDFEAFVSEADRDYICDRFPNLMVYVQFOHIMGLIPGVIIIS
YCIISKLSHSHKQKRAKTKTVIILILFFAKMLYYIGISDSTILEITIQGEF
ENIVHKWISTTEALAFHCCLNPLILYAFIAGKTKSAOHLIVSRSSSLIKLSKGR
GHSSVSTSESSSEFHS"
intron 1053. .3184
/gene="CXCR-4"
/number=1
exon 3185. .4760
/gene="CXCR-4"
/number=2
3'UTR 4229. .4760
/gene="CXCR-4"
polyA_signal 4726. .4731
/gene="CXCR-4"
BASE COUNT 1160 a 1338 c 1287 g 1376 t
ORIGIN

Query Match 90.0%; Score 1564; DB 11: Length 5161;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1564; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 ATATACACTTCAGATACTACACCGAGAAATGGGCTCAGGGGACTGACTCCATGAAG 165
DB 1185 ATATACACTTCAGATACTACACCGAGAAATGGGCTCAGGGGACTGACTCCATGAAG 3244
QY 156 GAACCGCTGTTCCGTGAGAAAATGCTAATTTCATATAAATCTTCTGCCACCACCTTAC 225
DB 3245 GAACCGCTGTTCCGTGAGAAAATGCTAATTTCATATAAATCTTCTGCCACCACCTTAC 3304
QY 226 TCCATCATCTTTTAAGTGCATTTGGGCATGATGGTGCATCCTGGTGCATGGGTAC 285
DB 3305 TCCATCATCTTTTAAGTGCATTTGGGCATGATGGTGCATCCTGGTGCATGGGTAC 3364
QY 286 CAGAGAACTGAGAGACATGACGACAGTACAGGCTGCACCTGTCACTGAGCCGACCTC 345
DB 3365 CAGAGAACTGAGAGACATGACGACAGTACAGGCTGCACCTGTCACTGAGCCGACCTC 3424
QY 346 CTCTTTGTCATCAGCGCTCCCTCTGGGCGATTGATCCGTGCAACTGGTACTTTGGG 405
DB 1185 CTCTTTGTCATCAGCGCTCCCTCTGGGCGATTGATCCGTGCAACTGGTACTTTGGG 405

DB 3425 CTCTTTGTCATCAGCGCTCCCTCTGGGCGATTGATCCGTGCAACTGGTACTTTGGG 3484
QY 406 AACTTCTTAATGCAAGGACATGCATGATCTACACAGTCAACCTCTACAGCAGTGTCTC 465
DB 3485 AACTTCTTAATGCAAGGACATGCATGATCTACACAGTCAACCTCTACAGCAGTGTCTC 3544
QY 466 ATCTGGCCTTCATCAGTGTGAGCGGTACCTGGCCATTCGTCACGCGCACCAACAGTCAG 525
DB 3545 ATCTGGCCTTCATCAGTGTGAGCGGTACCTGGCCATTCGTCACGCGCACCAACAGTCAG 3604
QY 526 AGGCCAAGAAAGCTGTGGCTGAAAAGTGGTATGTTGGCGTCTGGATCCCTGCCCTC 585
DB 3605 AGGCCAAGAAAGCTGTGGCTGAAAAGTGGTATGTTGGCGTCTGGATCCCTGCCCTC 3664
QY 586 CTGCTGACTATTCGGCAGTTCATCTTTGCCAAGCTCAGTGAAGCAGATGACAGTATATC 645
DB 3665 CTGCTGACTATTCGGCAGTTCATCTTTGCCAAGCTCAGTGAAGCAGATGACAGTATATC 3724
QY 646 TGTACCGCTTCACCCCAATGACTTGTGGGTGGTGTGTGTCAGTTTCAGACATCATG 705
DB 3725 TGTACCGCTTCACCCCAATGACTTGTGGGTGGTGTGTGTCAGTTTCAGACATCATG 3784
QY 706 GTTGGCCTTATCTGCTGCTGATTTGATCCTGTGCTGCTATTTGCATTAATCATCTCCAG 765
DB 3785 GTTGGCCTTATCTGCTGCTGATTTGATCCTGTGCTGCTATTTGCATTAATCATCTCCAG 3844
QY 766 CTGTCACATTCGCAAGGGCCACACGACGCAAGGCCCTCAAGACACAGTATCTCTATC 825
DB 3845 CTGTCACATTCGCAAGGGCCACACGACGCAAGGCCCTCAAGACACAGTATCTCTATC 3904
QY 826 CTGCTCTTCTGGCTGCTGCTGCTCTTACTACATTGGGATGACATGACTCCTTCATC 885
DB 3905 CTGCTCTTCTGGCTGCTGCTGCTCTTACTACATTGGGATGACATGACTCCTTCATC 3964
QY 886 CTCTGGAATTCATCAAGCAAGGCTGTGAGTTGAGAACCTGTGCACAGTGGATTTCC 945
DB 3965 CTCTGGAATTCATCAAGCAAGGCTGTGAGTTGAGAACCTGTGCACAGTGGATTTCC 4024
QY 946 ATCACGAGGCCCTAGGCTTCTCCATGTTGTGTGAAACCCATCTCTATGCTTCTT 1005
DB 4025 ATCACGAGGCCCTAGGCTTCTCTCCATGTTGTGTGAAACCCATCTCTATGCTTCTT 4084
QY 1006 GGAAGCCAAATTTAAACCTCTGCCACGACGACTCACTCTGTGAGCAGAGGCTCCAGC 1065
DB 4085 GGAAGCCAAATTTAAACCTCTGCCACGACGACTCACTCTGTGAGCAGAGGCTCCAGC 4144
QY 1066 CTCAGATCCTCTCCAAAGGAAAGCGAGGTGACATTCATCTGTTCCACTGAGTCTGAG 1125
DB 4145 CTCAGATCCTCTCCAAAGGAAAGCGAGGTGACATTCATCTGTTCCACTGAGTCTGAG 4204
QY 1126 TCTTCAAGTTTCACTCAGCTAACAACAGATGTAAGACATTTTATATACGATAATA 1185
DB 4205 TCTTCAAGTTTCACTCAGCTAACAACAGATGTAAGACATTTTATATACGATAATA 4264
QY 1186 ACTTTTTTTTAAAGTACACATTTTTCAGATATAAAGACTGACCAATATGTACAGTTTT 1245
DB 4265 ACTTTTTTTTAAAGTACACATTTTTCAGATATAAAGACTGACCAATATGTACAGTTTT 4324
QY 1246 TATTGCTGTGGATTTTGTCTTGTGTTCTTTAGTTTGTGAAATTAATTTGACTTA 1305
DB 4325 TATTGCTGTGGATTTTGTCTTGTGTTCTTTAGTTTGTGAAATTAATTTGACTTA 4384
QY 1306 TTTATATAAATTTTGTGTTGTTTCAATATGATGTGTCTGAGCAGAGCTGTGGCAAGT 1365
DB 4385 TTTATATAAATTTTGTGTTGTTTCAATATGATGTGTCTGAGCAGAGCTGTGGCAAGT 4444
QY 1366 TCTTAGTGTGTAATGTCTGTGTGAGAGCTGTAGAAAAGGAATGAAATTCACAGAGC 1425
DB 4445 TCTTAGTGTGTAATGTCTGTGTGAGAGCTGTAGAAAAGGAATGAAATTCACAGAGC 4504
QY 1426 GTGTAGGAATCAAGTAAAGCTGAAATGATCCCGAGCTGTTATGATGATTAATCTCT 1485
DB 4505 GTGTAGGAATCAAGTAAAGCTGAAATGATCCCGAGCTGTTATGATGATTAATCTCT 4564

```
QY 1486 CCATTCCCGTGAACGTTTTTCTCTTCTTAAGACGTGATTTGCTGTAGAACATGGCAC 1545
|||||
DB 4565 CCATTCCCGTGAACGTTTTTCTCTTCTTAAGACGTGATTTGCTGTAGAACATGGCAC 4624
|||||
QY 1546 TTATACCAAAAGCCCAAGGTGTATAGAAAATCGTGTTTTTCAGTTTTTCAGAGTGGGTT 1605
|||||
DB 4525 TTATACCAAAAGCCCAAGGTGTATAGAAAATCGTGTTTTTCAGTTTTTCAGAGTGGGTT 4684
|||||
QY 1606 GATTTCAGCACTCAGTGTACGTCTGTATTAAGTTGTTAATAAAGTACATGTTAAA 1665
|||||
DB 4685 GATTTCAGCACTCAGTGTACGTCTGTATTAAGTTGTTAATAAAGTACATGTTAAA 4744
|||||
QY 1666 CTTA 1669
|||||
DB 4745 CTTA 4748

RESULT 8
HSCXCR4 3733 bp DNA PRI 02-APR-1998
LOCUS Homo sapiens CXCR4 gene.
DEFINITION Y14739
ACCESSION G3021393
NID Y14739.1 GI:3021393
VERSION Chemokine receptor; CXCR4 gene.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 3733)
AUTHORS Friedl, R., Moepf, B. and Gierschik, P.
TITLE Genomic organization and expression pattern of the human chemokine
receptor CXCR4
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3733)
AUTHORS Friedl, R.
TITLE Direct Submission
JOURNAL Submitted (03-SEP-1997) R. Friedl, Department of Pharmacology and
Toxicology, Albert-Einstein-Allee 11, University of Ulm, 89081 Ulm,
FRG
REMARK revised by submitter 31-MAR-1998
FEATURES
Source location/Qualifiers
1..3733
/organism="Homo sapiens"
/db_xref="taxon:9606"
/rissue_type="blood"
join(<51..65,2198..>3241)
/gene="CXCR4"
<51..65
/gene="CXCR4"
/number=1
51..3241
/gene="CXCR4"
join(51..65,2198..3241)
/gene="CXCR4"
/codon_start=1
/product="chemokine receptor"
/protein_id="CAA75034.1"
/db_xref="PID:e1285584"
/db_xref="PID:g3021394"
/db_xref="GI:3021394"
/translation="MESISIYTSNTEEMSGSDYSMEKPECFRENNANFNKIFLPTI
YSLITLGIVNGLVTLVMGYOKLSMTDKYHLHLSVADLLFVITLPPWADVAVNM
YRGNPLCKAVHYVYNYLSSVILAFISLDRILAIVHINSORPKLLAEKVYVYG
WTPALLLTIPDIFAVNSADDRITCDRKYRNPIMVYVFOHIMVGLIPGIVISC
YCIILKLSHSGHOKRAKTVIILILAFRCMLPYIIGISIDSLILEIJKOGGEF
ENTVHKWISTTEALAFHCCCLNPLILVAFGAKFKTISOHRLIIVSNGSLKILSKGR
GSHSVSTSESSSFHSS"
65..2197
/gene="CXCR4"
/number=1
2198..>3241
intron
exon
```

```
BASE COUNT 822 a 955 c 980 g 976 t
ORIGIN
Query Match 88.4%; Score 1536; DB 10; Length 3733;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1536; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
/number=2
/gene="CXCR4"
QY 106 ATATACCTCAGTACTACACGAGAAATGGGCTCAGGGACATGACTCATGAAG 165
|||||
DB 2198 ATATACCTCAGTACTACACGAGAAATGGGCTCAGGGACATGACTCATGAAG 2257
|||||
QY 166 GAACCTGTTCCGTGAAGAAATGCTAATTTCAATAAAATCTTCTGCCACATCTAC 225
|||||
DB 2258 GAACCTGTTCCGTGAAGAAATGCTAATTTCAATAAAATCTTCTGCCACATCTAC 2317
|||||
QY 226 TCCATCATCTTCTTAAGTGGCATTTGGGCAATGGATTTGGTCACTCGTCATGGTTAC 285
|||||
DB 2318 TCCATCATCTTCTTAAGTGGCATTTGGGCAATGGATTTGGTCACTCGTCATGGTTAC 2377
|||||
QY 286 CAGAAAGAACTGAGAGATGACGACAGACAGTACAGCTGTGACGTGCGGACCTC 345
|||||
DB 2378 CAGAAAGAACTGAGAGATGACGACAGACAGTACAGCTGTGACGTGCGGACCTC 2437
|||||
QY 346 CTCTTTGTCATCAGGCTTCCCTTCTGGGCAATGATGCCGTGGCAAACTGTAATTTGGG 405
|||||
DB 2438 CTCTTTGTCATCAGGCTTCCCTTCTGGGCAATGATGCCGTGGCAAACTGTAATTTGGG 2497
|||||
QY 406 AACTTCCATGAGAGGCGATCATGTATCTACACAGCAACCTCTACAGAGTGTCTC 465
|||||
DB 2498 AACTTCCATGAGAGGCGATCATGTATCTACACAGCAACCTCTACAGAGTGTCTC 2557
|||||
QY 466 ATCTGGCCTTCATCATGCTGTGACCGCTACCTGGCCATGTCACAGCCACCAACAGTAG 525
|||||
DB 2558 ATCTGGCCTTCATCATGCTGTGACCGCTACCTGGCCATGTCACAGCCACCAACAGTAG 2617
|||||
QY 526 AGGCCAAGAACTGTTGGCTGAAAAAGTGCTATGTTGGCGTGTGATCCCTGCCCTC 585
|||||
DB 2618 AGGCCAAGAACTGTTGGCTGAAAAAGTGCTATGTTGGCGTGTGATCCCTGCCCTC 2677
|||||
QY 586 CTGTGACATATCCGACCTTCACTCTTCCCAACGTCACTGAGGCGAGATATATATC 645
|||||
DB 2678 CTGTGACATATCCGACCTTCACTCTTCCCAACGTCACTGAGGCGAGATATATATC 2737
|||||
QY 646 TGTGACCGCTTACCCCAATGACTTGTGGGTGTGTGTGTTCCATGTTACAGCACTATG 705
|||||
DB 2738 TGTGACCGCTTACCCCAATGACTTGTGGGTGTGTGTGTTCCATGTTACAGCACTATG 2797
|||||
QY 706 GTTGGCCTTATCTCGCTGGTATGTCATCCCTGCTGATTTGATTTATCTTCGAAG 765
|||||
DB 2798 GTTGGCCTTATCTCGCTGGTATGTCATCCCTGCTGATTTGATTTATCTTCGAAG 2857
|||||
QY 766 CTGTCACTTCCAGAGGCGCCAGAGCCAGAGGCCCTCAGACCACTATCCTCATC 825
|||||
DB 2858 CTGTCACTTCCAGAGGCGCCAGAGCCAGAGGCCCTCAGACCACTATCCTCATC 2917
|||||
QY 826 CTGGCTTTTGTGCTGTTGGCTGCTTACTACTATTTGGATAGAGATGACCTCTCATC 885
|||||
DB 2918 CTGGCTTTTGTGCTGTTGGCTGCTTACTACTATTTGGATAGAGATGACCTCTCATC 2977
|||||
QY 886 CTCTGGAATATCAAGAGAGGTGTGATTTGAGAACTGTGACAGTGGATTTCC 945
|||||
DB 2978 CTCTGGAATATCAAGAGAGGTGTGATTTGAGAACTGTGACAGTGGATTTCC 3037
|||||
QY 946 ATACAGAGGCCCTACTTCTTCCACTGTGTGTAACCCATCCTCTATGCTTTCTT 1005
|||||
DB 3038 ATACAGAGGCCCTACTTCTTCCACTGTGTGTAACCCATCCTCTATGCTTTCTT 3097
|||||
QY 1006 GAGCCAAATTTAAACCTCTGCCAGACGACGACCTGCTGTGAGAGAGGCTCAGC 1065
|||||
DB 3098 GAGCCAAATTTAAACCTCTGCCAGACGACGACGACCTGCTGTGAGAGAGGCTCAGC 3157
|||||
```



```
QY 466 ATCTGGCCCTTCATAGTGTGAGACCGCTACCTGGCCATGCTCCAGCCGACCAAGACTGAG 525
    |||
DB 5228 ATTCCTGGCCCTTCATAGTGTGAGACCGCTACCTGGCCATGCTCCAGCCGACCAAGACTGAG 5287
QY 526 AGCCCAAGAGAGCTGTGGCTGAAAAGGTGTATGTGGCTGTGAGATCCCTGCCCTC 585
    |||
DB 5288 AGCCCAAGAGAGCTGTGGCTGAAAAGGTGTATGTGGCTGTGAGATCCCTGCCCTC 5347
QY 586 CTGCTGACTATTCGCCGACTTCATCTTTGCCAGGTGAGAGGAGAGATGACATATATC 645
    |||
DB 5348 CTGCTGACTATTCGCCGACTTCATCTTTGCCAGGTGAGAGGAGAGATGACATATATC 5407
QY 646 TGTGACCGCTTCACCCCAATGACTGTGGGTGGTGTGTCAGATTTCAGCATCATG 705
    |||
DB 5408 TGTGACCGCTTCACCCCAATGACTGTGGGTGGTGTGTCAGATTTCAGCATCATG 5467
QY 706 GTTGGGCTTATCCGCTGTGATGTGATGCTGCTGCTATTTGCAATATCATCTCCAG 765
    |||
DB 5458 GTTGGGCTTATCCGCTGTGATGTGATGCTGCTGCTATTTGCAATATCATCTCCAG 5527
QY 766 CTGTCACTCTCCAAAGGCGCACAGAGCGCAAGGCCCTCAAGACCAAGTATCTCATC 825
    |||
DB 5528 CTGTCACTCTCCAAAGGCGCACAGAGCGCAAGGCCCTCAAGACCAAGTATCTCATC 5587
QY 826 CTGCTTTCTTCCGCTGTGGCTGCTTACTACATGAGATGAGATGAGATGAGATGAGAT 885
    |||
DB 5588 CTGCTTTCTTCCGCTGTGGCTGCTTACTACATGAGATGAGATGAGATGAGATGAGAT 5647
QY 886 CTCTGGAATATCATCAAGAGAGGTGTGAGTTTGAACACATGTGCACAGTGTGATTC 945
    |||
DB 5648 CTCTGGAATATCATCAAGAGAGGTGTGAGTTTGAACACATGTGCACAGTGTGATTC 5707
QY 946 ATCACCAGAGCCCTAGCTTTCTTCCACGTGTGTGAACCCCATCTCTATGCTTCTT 1005
    |||
DB 5708 ATCACCAGAGCCCTAGCTTTCTTCCACGTGTGTGAACCCCATCTCTATGCTTCTT 5767
QY 1006 GGAGCCCAATTTAAAGCTGTGCCAGACGACACTCCTGTGTGAGAGAGAGGTGTCAGC 1065
    |||
DB 5768 GGAGCCCAATTTAAAGCTGTGCCAGACGACACTCCTGTGTGAGAGAGAGGTGTCAGC 5827
QY 1066 CTCAAGATCTCTCCAAAGGAGAGGAGTGTGACATTTATCTGTTTCCACGTGAGTCTAG 1125
    |||
DB 5828 CTCAAGATCTCTCCAAAGGAGAGGAGTGTGACATTTATCTGTTTCCACGTGAGTCTAG 5887
QY 1126 TCTTCAAGTCTTCACTCCAGCTAACACAGATGTAAAGACTTTTATACGATTAATA 1185
    |||
DB 5888 TCTTCAAGTCTTCACTCCAGCTAACACAGATGTAAAGACTTTTATACGATTAATA 5947
QY 1186 ACTTTTTTAAAGTATACATTTTTCAGATATATAAGACTGACCAATATTTGACGTTT 1245
    |||
DB 5948 ACTTTTTTAAAGTATACATTTTTCAGATATATAAGACTGACCAATATTTGACGTTT 6007
QY 1246 TATTCCTGTGGAGATTTTGTCTTGTGTTTCTTATGATTTTGTGAAGTTAATGACTTA 1305
    |||
DB 6008 TATTCCTGTGGAGATTTTGTCTTGTGTTTCTTATGATTTTGTGAAGTTAATGACTTA 6067
QY 1306 TTTATATAATTTTTTTTGTTCATATATGATGTGTGTAGCAGAGACCTGTGGCAAGT 1365
    |||
DB 6068 TTTATATAATTTTTTTTGTTCATATATGATGTGTGTAGCAGAGACCTGTGGCAAGT 6127
QY 1366 TCTTAGTTGCTGTATGTCGTGTGTAGAGACTGTAGAAAAGGAACTGAACATTTCCAGAGC 1425
    |||
DB 6128 TCTTAGTTGCTGTATGTCGTGTGTAGAGACTGTAGAAAAGGAACTGAACATTTCCAGAGC 6187
QY 1426 GTGTAGTGAATCAAGTAAGCTAGAAATGATCCCAAGCTGTTTATGATAGATATCTCT 1485
    |||
DB 6188 GTGTAGTGAATCAAGTAAGCTAGAAATGATCCCAAGCTGTTTATGATAGATATCTCT 6247
QY 1486 CCATTCCTCGGAGAGCTTTTCTGTTCTTAAGAGCTGATTTTCTGTAGAGATGGCAC 1545
    |||
DB 6248 CCATTCCTCGGAGAGCTTTTCTGTTCTTAAGAGCTGATTTTCTGTAGAGATGGCAC 6307
QY 1546 TTATAACCAAAAGCCAAAGTGTATAGAAATGCTGTTTTCAGATTTCAGGAGTGGGT 1605
```

```
DB 6308 TTATAACCAAAAGCCAAAGTGTATAGAAATGCTGTTTTCAGATTTCAGGAGTGGGT 6367
QY 1606 GATTTCAGACCTACAGTGTAGAGCTTGTATTAAGTTTAAAGTACATGTAAA 1665
    |||
DB 6368 GATTTCAGACCTACAGTGTAGAGCTTGTATTAAGTTTAAAGTACATGTAAA 6427
QY 1666 CTTA 1669
    |||
DB 6428 CTTA 6431

RESULT 10
AF052572 8747 bp DNA PRI 07-SEP-1998
LOCUS Homo sapiens chemokine receptor CXCR4 gene, promoter region and
DEFINITION complete cds.
ACCESSION AF052572
NID 93549254
VERSION AF052572.1 GI:3549254
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 8747)
            Caruz,A., Samsom,M., Alonso,J.M., Alcamí,J., Balleux,F.,
            Virelizier,J.L., Parmentier,M., and Arenzana-Seisdedos,F.
            Genomic organization and promoter characterization of human CXCR4
            gene.
            FEBS Lett. 426 (2), 271-278 (1998)
JOURNAL 96258970
MEDLINE 2 (bases 1 to 8747)
REFERENCE Caruz,A., Samsom,M., Virelizier,J.L., Parmentier,M. and
AUTHORS Arenzana-Seisdedos,F.
TITLE Direct Submission
JOURNAL Submitted (06-MAR-1998) Immunologie Virale, Institut Pasteur, 28
Rue Dr Roux, Paris 75724, France
FEATURES
            Location/Qualifiers
            1..8747
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="24D3"
            /note="BAC clone from Research Genetics"
            1..2642
            /join(2643..2735,4868..6429)
            /product="chemokine receptor CXCR4"
            /join(2721..2735,4868..5911)
            /codon_start=1
            /product="chemokine receptor CXCR4"
            /protein_id="AAC34581.1"
            /db_xref="PIR:g3549255"
            /db_xref="PIR:g3549255"
            /translation="MEGISITSDNYTEMGSGDYDSKPEPRENANFNKFLPTI
            YSILFLGIVNGVLIVMGYOKRLRSMTRKYRLHSVADLLFVITLPEAVDAVANM
            YFEGFLKRAVAVITVNLVSSVLLAFISDRYLAIVAHNSQPRRLAEKVAVG
            WIPALLILPEIPFANVSEADDRYICDRFENPDLAVVFOHIMVLIIPGIVLSC
            YCIITSLSHSKGRKRAKTKTYILIAFACALPYIGISDSFLLIEIKGCEFE
            ENYHKMISTIEALFHCINPLIYVFLAKFKTSMQHALTSVRSSSLKLSKGR
            GHSVSSTSESSSFHSS"
BASE COUNT 2161 a 2173 c 1966 g 2447 t
ORIGIN
Query Match 87.1% Score 1513: DB 11: Length 8747:
Best Local Similarity 99.9%: Fred. No. 0:
Matches 1563: Conservative 0: Mismatches 1: Indels 0: Gaps 0:
```

```
|||||
Db 4928 GAACCCGTTCGTCAGAAATGCTAATTTCAATAAAATCTTCCGCCACATCTAC 4987
QY 225 TCCATCATCTTTAACTGGCAATTTGGCCAAATGATGGTCACTCTGGTCATAGGTTAC 285
Db 4988 TCCATCATCTTTAACTGGCAATTTGGCCAAATGATGGTCACTCTGGTCATAGGTTAC 5047
QY 285 CAGAAAGAACTGAGAGCATGACGAGCAAGTACAGGTGACCTGTGAGGGCGAGCTC 345
Db 5048 CAGAAAGAACTGAGAGCATGACGAGCAAGTACAGGTGACCTGTGAGGGCGAGCTC 5107
QY 345 CTCCTTTGATCATAGCCTTCCCTTCTGGGCAAGTATGATGCCGTGGCAACTGGTACTTTGGG 405
Db 5108 CTCCTTTGATCATAGCCTTCCCTTCTGGGCAAGTATGATGCCGTGGCAACTGGTACTTTGGG 5167
QY 405 AACTTCTTATGCAAGGCAATGCTCATCTACAGTCAAGTCAACTCTACAGCAATGCTC 465
Db 5168 AACTTCTTATGCAAGGCAATGCTCATCTACAGTCAAGTCAACTCTACAGCAATGCTC 5227
QY 465 ATCTTGCCCTTATCATAGTCTGACCGCTACCTGACCTGCTCCAGCCCAACAGTACAG 525
Db 5228 ATCTTGCCCTTATCATAGTCTGACCGCTACCTGACCTGCTCCAGCCCAACAGTACAG 5287
QY 525 AGCCCAAGCAAGCTGTGGCTGAAGAGTGCTATGTTGGCGTCTGGATCCTGCCCTC 585
Db 5288 AGCCCAAGCAAGCTGTGGCTGAAGAGTGCTATGTTGGCGTCTGGATCCTGCCCTC 5347
QY 585 CTGCTGACTATTCGCGACTTCATCTTGCCCAAGTCAAGTCAAGTCAAGTCAAGTATATC 645
Db 5348 CTGCTGACTATTCGCGACTTCATCTTGCCCAAGTCAAGTCAAGTCAAGTCAAGTATATC 5407
QY 645 TGTGACCGCTTCTACCCCAATGACTTGTGGGTGTTGTGTCCAGTTTACGACATCATG 705
Db 5408 TGTGACCGCTTCTACCCCAATGACTTGTGGGTGTTGTGTCCAGTTTACGACATCATG 5467
QY 705 GTTGGCTTATCTGCTGCTGTGATGTGATCTCTCTGCTATGATATCATCTCCAG 765
Db 5468 GTTGGCTTATCTGCTGCTGTGATGTGATCTCTCTGCTATGATATCATCTCCAG 5527
QY 765 CTGTCACTATCCAGAGGCGCCAGAGGCGCCCAAGGCGCCCAAGGCGCCCAAGGCGCCCAAG 825
Db 5528 CTGTCACTATCCAGAGGCGCCAGAGGCGCCCAAGGCGCCCAAGGCGCCCAAGGCGCCCAAG 5587
QY 825 CTGGCTTTTGGCTGTGGCTGCTTACTACATTTGGATGAGATCGACTCCTTCATC 885
Db 5588 CTGGCTTTTGGCTGTGGCTGCTTACTACATTTGGATGAGATCGACTCCTTCATC 5647
QY 885 CTCTCGAATATCATCAGAGGCGGTGATTTGAGAACACTGTGCAAGAGTGGATTTC 945
Db 5648 CTCTCGAATATCATCAGAGGCGGTGATTTGAGAACACTGTGCAAGAGTGGATTTC 5707
QY 945 ATCAGCGAGGCGCTAGCTTTCTTCCACTGTGTCTGAACCCCACTCTATGCTTCTCT 1005
Db 5708 ATCAGCGAGGCGCTAGCTTTCTTCCACTGTGTCTGAACCCCACTCTATGCTTCTCT 5767
QY 1005 GGAGCCAAATTTAAACCTCTGCGCCAGCAGCACTACCTCTGTGAGAGGCGTCCAG 1065
Db 5768 GGAGCCAAATTTAAACCTCTGCGCCAGCAGCACTACCTCTGTGAGAGGCGTCCAG 5827
QY 1065 CTCAGAGTCTCTCCAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1125
Db 5828 CTCAGAGTCTCTCCAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5887
QY 1125 TCTTCAAGTTTCACTCAGCTAACAAGATGTAAAGACTTTTATAGCATTAATA 1185
Db 5888 TCTTCAAGTTTCACTCAGCTAACAAGATGTAAAGACTTTTATAGCATTAATA 5947
QY 1186 ACTTTTATAGTATACATTTTATAGATATATAAGACTGACCATATTTGATCACTTTT 1245
Db 5948 ACTTTTATAGTATACATTTTATAGATATATAAGACTGACCATATTTGATCACTTTT 6007
QY 1245 TATTGCTTGTGATTTTGTGTTCTTCTTAACTTTTGTGAAGTTAATGACTTA 1305
|||||
```

```
Db 6008 TATTGCTTGTGATTTTGTGTTCTTCTTAACTTTTGTGAAGTTAATGACTTA 6067
QY 1306 TTTATATAAATTTTGTGTTTATTTGATATTTGATGTGTCTAGCGAGCACTGTGGCCAAGT 1365
Db 6068 TTTATATAAATTTTGTGTTTATTTGATATTTGATGTGTCTAGCGAGCACTGTGGCCAAGT 6127
QY 1366 TCTTATGCTGTATGCTGTGTGAGTGTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1425
Db 6128 TCTTATGCTGTATGCTGTGTGAGTGTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6187
QY 1426 GTGTAGTGAATCAGTAAAGCTAGAAATGATCCGAGCTGTTTATGATAGATATCTCT 1485
Db 6188 GTGTAGTGAATCAGTAAAGCTAGAAATGATCCGAGCTGTTTATGATAGATATCTCT 6247
QY 1486 CCAATCCCGTGGAGAGCTTTTCTGTTTAAAGCTGATTTTGTGTGTAAGATGGCAC 1545
Db 6248 CCAATCCCGTGGAGAGCTTTTCTGTTTAAAGCTGATTTTGTGTGTAAGATGGCAC 6307
QY 1546 TTATAACCAAGGCCAAAGTGTATAGAAATGCTGTTTTCAGTTTTCAGAGTGGGTT 1605
Db 6308 TTATAACCAAGGCCAAAGTGTATAGAAATGCTGTTTTCAGTTTTCAGAGTGGGTT 6367
QY 1606 GATTTCAGCACTACAGTGTACAGTCTTGTATTAAGTGTATAAAGTACATGTTAAA 1665
Db 6368 GATTTCAGCACTACAGTGTACAGTCTTGTATTAAGTGTATAAAGTACATGTTAAA 6427
QY 1665 CTTA 1669
Db 6428 CTTA 6431

RESULT 11
HUMSTR 1637 bp mRNA PRI 10-MAR-1994
LOCUS Homo sapiens seven transmembrane segment receptor mRNA, complete
DEFINITION
ACCESSION M99293
NID 9292516
VERSION M99293.1 GI:292516
KEYWORDS seven transmembrane segment receptor.
SOURCE Homo sapiens (human).
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1637)
Federspiel, B., Melhado, I. G., Duncan, A. M., Delaney, A.,
Schappert, K., Clark-Lewis, I., and Jirik, F. R.
Molecular cloning of the cDNA and chromosomal localization of the
gene for a putative seven-transmembrane segment (7-TMS) receptor
isolated from human spleen
JOURNAL Molecular cloning of the cDNA and chromosomal localization of the
MEDLINE isolated from human spleen
FEATURES
source 93315164
misc_feature location:Qualifiers
1..1637
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="2q21"
1
/feature="GDB identifier D25201E; S9a localized by in situ
hybridization to human chromosome 2q21"
85..1143
/note="93% homology to bovine neuropeptide Y3; precursor;
putative"
/codon_start=1
/product="seven transmembrane segment receptor"
/protein_id="AA16617.1"
/db_xref="PDB:9292517"
/db_xref="GI:292517"
/translation="MEGISIVTSDNYEEMSGSDYSKPECFREENANFKILPFTI
YSILFLGIVNGAVLIVMGYQKRLRMTDKYRLHLVADLLFVITLPEFAVDVAVM
YGFNLCRAVHIVTVNLVSYLLAIFSLDRYLALVHATNSQRPRLLAERVVYGV
WIPALLETIPFIVANVEADRDYICDRFYNDLWVYVFOFOHIVGILPLGIYLSG
CIIISKSHSKGKRAKLTIVILLIAFACWLPYIGISIDSFILLEIKQGER
```


Best Local Similarity 99.8%; Pred. No. 0;
Matches 1555; Conservative 0; Mismatches 1; Indels 3; Gaps 3;

```

QY 30 AAGTACGCCGAGGCTTGAGTGTCCAGTAGCCACCGCATCTGGAGAACCGGGTTAC 89
    |||
Db 12 AAGTACGCCGAGGCTTGAGTGTCCAGTAGCCACCGCATCTGGAGAACCGGGTTAC 71
    |||
QY 90 CATGGAGGGGATGATATACATTCAGATTAACATACCGAGGAATGGGCTAGGGGA 149
    |||
Db 72 CATGGAGGGGATGATATACATTCAGATTAACATACCGAGGAATGGGCTAGGGGA 131
    |||
QY 150 CTATGATCCATGAAGAACCCCTGTTCCGTGAAGAAATGCTAATTTCATAAATCTT 209
    |||
Db 132 CTATGATCCATGAAGAACCCCTGTTCCGTGAAGAAATGCTAATTTCATAAATCTT 191
    |||
QY 210 CCGGCCACCATCTACTCCATCTTCTTACATGCGCATGTGGGCAATGGATTGTCAT 269
    |||
Db 192 CCGGCCACCATCTACTCCATCTTCTTACATGCGCATGTGGGCAATGGATTGTCAT 251
    |||
QY 270 CCTGTGATGGGTTACGAGAGAAACTGAGAGCATGACGGACAGTACAGGCTGACCT 329
    |||
Db 252 CCGTGTATGGGTTACGAGAGAAACTGAGAGCATGACGGACAGTACAGGCTGACCT 311
    |||
QY 330 GTCATGCGCCGACCTCTCTTTGTATCAAGCTTCCCTTGTGGCAGTTGATGCCGTGGC 389
    |||
Db 312 GTCATGCGCCGACCTCTCTTTGTATCAAGCTTCCCTTGTGGCAGTTGATGCCGTGGC 371
    |||
QY 390 AACTGTGATTTGGGAGCTTCATGCAAGGAGTGCATGATCTACACAGTCAACT 449
    |||
Db 372 AACTGTGATTTGGGAGCTTCATGCAAGGAGTGCATGATCTACACAGTCAACT 431
    |||
QY 450 CTACAGCAGTGTCTCATCTGCGCTTCATCAAGTGTGACCGCTACCTGGCCATCGTCCA 509
    |||
Db 432 CTACAGCAGTGTCTCATCTGCGCTTCATCAAGTGTGACCGCTACCTGGCCATCGTCCA 491
    |||
QY 510 CCGCCACCAACAGTCAGAGGCCAAGGAAGCTGTGGCTGAAGAAAGTGCTATGTGGCGT 569
    |||
Db 492 CCGCCACCAACAGTCAGAGGCCAAGGAAGCTGTGGCTGAAGAAAGTGCTATGTGGCGT 551
    |||
QY 570 CTGATCCCTGCGCTCTGCTGACATATCCGAGCTTCATCTTGGCCAGCTCAGTGAAGC 629
    |||
Db 552 CTGATCCCTGCGCTCTGCTGACATATCCGAGCTTCATCTTGGCCAGCTCAGTGAAGC 611
    |||
QY 630 AGATGACAGATATATCTGTGACCGCTTCTACCCCAATGACTTGTGGTGTGTGTTCCA 689
    |||
Db 612 AGATGACAGATATATCTGTGACCGCTTCTACCCCAATGACTTGTGGTGTGTGTTCCA 671
    |||
QY 690 GTTTCAGCATCATCTGTTGGCTTATCCCTGCTGTGATTTGTATCTGCTCCGTATTTG 749
    |||
Db 672 GTTTCAGCATCATCTGTTGGCTTATCCCTGCTGTGATTTGTATCTGCTCCGTATTTG 731
    |||
QY 750 CATTATCATCTCAACCTGTCACTCTCAAGGCCACACAGGCGCAAGGCCCTCAAGAC 809
    |||
Db 732 CATTATCATCTCAACCTGTCACTCTCAAGGCCACACAGGCGCAAGGCCCTCAAGAC 791
    |||
QY 810 CACAGTCATCTCATCTGCGCTTCTTCCGCTGTGGCTGTGCTTACTACATTGGATTCAG 869
    |||
Db 792 CACAGTCATCTCATCTGCGCTTCTTCCGCTGTGGCTGTGCTTACTACATTGGATTCAG 851
    |||
QY 870 CATGCGCTCTCATCTCTCTGGAATTCATCAAGCAAGGTTGACTTTGAGAACCTGT 929
    |||
Db 852 CATGCGCTCTCATCTCTCTGGAATTCATCAAGCAAGGTTGACTTTGAGAACCTGT 911
    |||
QY 930 GCACAAGTGAATTTCCATCAGCAGGCGCTAGCTTCTTCCACTGTTGTGTGAACCCAT 989
    |||
Db 912 GCACAAGTGAATTTCCATCAGCAGGCGCTAGCTTCTTCCACTGTTGTGTGAACCCAT 971
    |||
QY 990 CCTCTATGCTTTCTTGGAGCCAAATTTAAACCTGTGCCACAGCAGCTACCTCTGT 1049
    |||
Db 972 CCTCTATGCTTTCTTGGAGCCAAATTTAAACCTGTGCCACAGCAGCTACCTCTGT 1031
    |||
QY 1050 GAGCAGAGGGTCCAGCCTCAAGATCTCTCAAGGAAAGCGAGTGGACATTCATCTGT 1109
    |||

```

```

Db 1032 GAGCAGAGGGTCCAGCCTCAAGATCCTCTCCAAAGGAAGGAGTGGACATTCATCTGT 1091
QY 1110 TTCACCTAGCTGTGAGTCTTCAAGTTTTCACCTCCAGCTAACACAGATGTAAGAAGCTTTT 1169
    |||
Db 1092 TTCACCTAGCTGTGAGTCTTCAAGTTTTCACCTCCAGCTAACACAGATGTAAGAAGCTTTT 1151
    |||
QY 1170 TTTTATGCAATTAATACCTTTTATAGTTACACATTTTTCAGATTAATAAGAGTACC 1229
    |||
Db 1152 TTTTATGCAATTAATACCTTTTATAGTTACACATTTTTCAGATTAATAAGAGTACC 1211
    |||
QY 1230 AATATTGTACAGTTTATATGCTTGTGGATTTTTCGTTGTTCTTCTTACTTGT 1287
    |||
Db 1212 AATATTGTACAGTTTATATGCTTGTGGATTTTTCGTTGTTCTTCTTACTTGT 1271
    |||
QY 1288 TGAATGTTTATGACTTATTTATTAATTTTGTGTTTTCATATGATGTGTCTAG 1346
    |||
Db 1272 TGAAGGTTTATGACTTATTTATTAATTTTGTGTTTTCATATGATGTGTCTAG 1331
    |||
QY 1347 GCAGAGCCTGTGGCCAAAGTCTTGTGCTGTATGCTGTGGTGTAGGACTGTGAAGAAAG 1406
    |||
Db 1332 GCAGAGCCTGTGGCCAAAGTCTTGTGCTGTATGCTGTGGTGTAGGACTGTGAAGAAAG 1391
    |||
QY 1407 GAACGTGAACATTCACAGAGCGTGTAGTAATCAGCTAAAGCTAGAAATGATCCCAAGCTGT 1466
    |||
Db 1392 GAACGTGAACATTCACAGAGCGTGTAGTAATCAGCTAAAGCTAGAAATGATCCCAAGCTGT 1451
    |||
QY 1467 TTATGCTATGATATCTCTTCATCTCCGCTGGAGACGTTTTTCTGTTCTTAAAGCGTAT 1526
    |||
Db 1452 TTATGCTATGATATCTCTTCATCTCCGCTGGAGACGTTTTTCTGTTCTTAAAGCGTAT 1511
    |||
QY 1527 TTGCTGTAGAAGATGGACACTTATTAACCAAGGCCCAAGTGTATAGAATGCTGTGTTT 1586
    |||
Db 1512 TTGCTGTAGAAGATGGACACTTATTAACCAAGGCCCAAGTGTATAGAATGCTGTGTTT 1571
    |||
QY 1587 CAGTTTTCAGAGATGGTGTATTTTCAGACCTAGACAGTGTGATTTGATTAAGTTGT 1646
    |||
Db 1572 CAGTTTTCAGAGATGGTGTATTTTCAGACCTAGACAGTGTGATTTGATTAAGTTGT 1631
    |||
QY 1647 AATTAAGTACATGTTAACTTAAAAA 1685
    |||
Db 1632 AATTAAGTACATGTTAACTTAAAAA 1670
    |||

```

RESULT 13
 HUMNYRECA
 LOCUS
 DEFINITION
 ACCESSION
 NID
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1225)
 Jazlin,E.E., Yeo,H., Blomqvist,A.G., Yee,F., Weng,G., Walker,M.W.,
 Salton,J., Larnhammar,D. and Wahlestedt,C.
 A proposed bovine neuropeptide Y (NPY) receptor cDNA clone, or its
 human homologue, confers neither NPY binding sites nor NPY
 responsiveness on transfected cells
 Regul. Pept. 47 (3), 247-258 (1993)
 94052833
 FEATURES
 MEDLINE
 JOURNAL
 TITLE
 AUTHORS
 REFERENCE
 ORGANISM
 SOURCE
 Location/Qualifiers
 1..1225
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 26..1084
 /gene="NPYR"
 26..1084
 /gene="NPYR"

```

/note="putative"
/codon_start=1
/product="neuropeptide Y receptor"
/protein_id="AA16594.1"
/db_xref="PID:q189314"
/db_xref="GI:189314"
/translation="MEGISITVSDNTYEMSGDYDSMKPECFREANFNKILPTI
YSIIFLTGIVNGLVILVMGYOKLRSMIDKRLHSVADLLFVITLPPWADVAMN
YFENFLCAVHYIYVNLVSVLLIAFISLDRLAIYHATNSQRPRLAEKVYVGV
WIPALLTIIDPFIANVSEADRYICDFPNIDLVVFOHIMVGLIPGIYILSC
YCIISKLSHSGHOKRAKLTIVILIAFPACWLPYIGISIDSIILEIYIKGCE
ENTVHKWISITELAFHCCLPILYAFGAKFKTSQAHALISVSGSSKLILSKGR
GGHSVSTSESSSFHSS"
BASE COUNT      288 a      324 c      259 g      354 t
ORIGIN
Query Match      63.8%; Score 1108; DB 9; Length 1225;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 56 CGCATCTGAGAACCCAGCGGTACATGAGGGGATGAGTATATACCTTCAGATACTA 125
    |||
Db 1 CGCATCTGAGAACCCAGCGGTACATGAGGGGATGAGTATATACCTTCAGATACTA 60
QY 126 CACCGAGGAATGGGCTCAGGGGACTATGACTCCATGAGGAACCCCTGTTCCGTGA 185
    |||
Db 51 CACCGAGGAATGGGCTCAGGGGACTATGACTCCATGAGGAACCCCTGTTCCGTGA 120
QY 186 AATGCTAATTTCAATTAATCTTCCGCCACCATCTACTCCATCTTTCTTAAGTG 245
    |||
Db 121 AATGCTAATTTCAATTAATCTTCCGCCACCATCTACTCCATCTTTCTTAAGTG 180
QY 246 CATTTGGCGAATGATTTGATCTCTGTCATGAGGTTTACAGAGAAGTGAAGACAT 305
    |||
Db 181 CATTTGGCGAATGATTTGATCTCTGTCATGAGGTTTACAGAGAAGTGAAGACAT 240
QY 306 GACGAGCAATAGAGGCTGACCTGTGAGTGGCCGACCTCTCTTTGTATCAGCCTTC 365
    |||
Db 241 GACGAGCAATAGAGGCTGACCTGTGAGTGGCCGACCTCTCTTTGTATCAGCCTTC 300
QY 366 CTCTCTGGGAGTTGATCTCGGTGGCAAACTGGTACTTTGGGAACTCTTATGCAAGG 425
    |||
Db 301 CTCTCTGGGAGTTGATCTCGGTGGCAAACTGGTACTTTGGGAACTCTTATGCAAGG 360
QY 426 CCATGTATATACAGTCAACCTCTACAGAGAGTCCATCTGCTGACCTTCATCACTCT 485
    |||
Db 361 CCATGTATATACAGTCAACCTCTACAGAGAGTCCATCTGCTGACCTTCATCACTCT 420
QY 486 GGACCGCTACCTGGCCATCTGCCAGCCACCAACAGTCAAGGCCAAGAGCTGTTGGC 545
    |||
Db 421 GGACCGCTACCTGGCCATCTGCCAGCCACCAACAGTCAAGGCCAAGAGCTGTTGGC 480
QY 546 TGAAGAAGTGTCTATGTTGGCGTCTGGATCCCGCCTCTCTGTGACTATTCGCCACTT 605
    |||
Db 481 TGAAGAAGTGTCTATGTTGGCGTCTGGATCCCGCCTCTCTGTGACTATTCGCCACTT 540
QY 606 CATCTTTGCCAAGCTGACGTAGGAGAGATGACATATATCTGTGACCGCTTCACCCCA 665
    |||
Db 541 CATCTTTGCCAAGCTGACGTAGGAGAGATGACATATATCTGTGACCGCTTCACCCCA 600
QY 666 TGACTTGTGGGTGGTGTGTTCCAGTTTACAGACATGAGTTGGCTTATCTGCTGG 725
    |||
Db 601 TGACTTGTGGGTGGTGTGTTCCAGTTTACAGACATGAGTTGGCTTATCTGCTGG 660
QY 726 TATTTGATCTCTGCTGCTGATATGATATATCTTCAAGCTGTCAACCTCCAAAGGCCA 785
    |||
Db 661 TATTTGATCTCTGCTGCTGATATGATATATCTTCAAGCTGTCAACCTCCAAAGGCCA 720
QY 786 CCGAAGACGAGGCGCTCAAGACACAGTATCTGATCTGAGCTTTCTGCGCTGTG 845
    |||
Db 721 CCGAAGACGAGGCGCTCAAGACACAGTATCTGATCTGAGCTTTCTGCGCTGTG 780

```

```

QY 846 GCTGCCCTTACTACATTTGGGATCAGATCGATCTCTTATCTCTCTGGAATCATCAAGA 905
    |||
Db 781 GCTGCCCTTACTACATTTGGGATCAGATCGATCTCTTATCTCTCTGGAATCATCAAGA 840
QY 906 AGGCTGTGATTTGGAACACGTGTGACAGAGTGGATTTCCATCACCAGGCGCTTGCCTT 965
    |||
Db 841 AGGCTGTGATTTGGAACACGTGTGACAGAGTGGATTTCCATCACCAGGCGCTTGCCTT 900
QY 966 CTTCACATGTTGCTGGAACCCCATCTCTATGCTTCTCTGAGGCCAATTTAAACCTC 1025
    |||
Db 901 CTTCACATGTTGCTGGAACCCCATCTCTATGCTTCTCTGAGGCCAATTTAAACCTC 960
QY 1026 TGCCGACGACGACCTACCTGTGTGACAGAGAGGTTCACGCTCAAGATCTCTCCAAG 1085
    |||
Db 961 TGCCGACGACGACCTACCTGTGTGACAGAGAGGTTCACGCTCAAGATCTCTCCAAG 1020
QY 1086 AAGCGAGGTGACATTCATCTGTTCCATGAGTCTGAGTCTCAAGTTTCACTCCAG 1145
    |||
Db 1021 AAGCGAGGTGACATTCATCTGTTCCATGAGTCTGAGTCTCAAGTTTCACTCCAG 1080
QY 1146 CTACACAGATGTAAAGACTTTTCTT 1173
    |||
Db 1081 CTACACAGATGTAAAGACTTTTCTT 1108
RESULT 14
AF025375
LOCUS AF025375 1059 bp mRNA PRI 28-OCT-1997
DEFINITION Homo sapiens chemokine receptor-4 (CXCR4) mRNA, complete cds.
ACCESSION AF025375
NID 92565335
VERSION AF025375.1 GI:2565335
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1
Xiao,L., Weiss,S., Qari,S., Rudolph,D., Hodge,T. and Lal,R.
Partial resistance to infection by syncytium-inducing primary HIV-1
in exposed uninfected individuals homozygous for CCR5 32bp deletion
Unpublished
2 (bases 1 to 1059)
JOURNAL Qari,S.
AUTHORS Direct Submission
TITLE Submitted (17-SEP-1997) Retrovirus Diseases Branch, Centers for
Diseases Control and Prevention, 1600 Clifton Road, Atlanta, GA
30333, USA
FEATURES
source location/Qualifiers
1..1059
/organism="Homo sapiens"
/isolate="patient EUD32-1"
/db_xref="taxon:9606"
/chromosome="2"
/map="2q21"
1..1059
/gene="CXCR4"
1..1059
/feature="CXCR4"
/note="fusin: G protein-coupled receptor; HUMSTR: LESTR"
/codon_start=1
/product="chemokine receptor-4"
/protein_id="AAB81970.1"
/db_xref="PID:92565336"
/db_xref="GI:2565336"
/translation="MEGISITVSDNTYEMSGDYDSMKPECFREANFNKILPTI
YSIIFLTGIVNGLVILVMGYOKLRSMIDKRLHSVADLLFVITLPPWADVAMN
YFENFLCAVHYIYVNLVSVLLIAFISLDRLAIYHATNSQRPRLAEKVYVGV
WIPALLTIIDPFIANVSEADRYICDFPNIDLVVFOHIMVGLIPGIYILSC
YCIISKLSHSGHOKRAKLTIVILIAFPACWLPYIGISIDSIILEIYIKGCE
ENTVHKWISITELAFHCCLPILYAFGAKFKTSQAHALISVSGSSKLILSKGR
GGHSVSTSESSSFHSS"
BASE COUNT      244 a      301 c      232 g      282 t

```

ORIGIN

Query Match 61.0%; Score 1059; DB 11; Length 1059;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1059; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 91 ATGGAGGGGATCGATATATACACTCAGATACATACACCGAGAAATGGGCTCAGGGAC 150
    |||||||
Db 1 ATGGAGGGGATCGATATATACACTCAGATACATACACCGAGAAATGGGCTCAGGGAC 60
    |||||||
QY 151 TATGACTCAGTGAAGAACCCCTGTTCCGTGAAGAAAAATCTAATTTCAATAAATCTTC 210
    |||||||
Db 51 TATGACTCAGTGAAGAACCCCTGTTCCGTGAAGAAAAATCTAATTTCAATAAATCTTC 120
    |||||||
QY 211 CTGCCCCACCATCTACTCAGATCTCTTACTGCGATTTGTGGCAATGGAATGGTATC 270
    |||||||
Db 121 CTGCCCCACCATCTACTCAGATCTCTTACTGCGATTTGTGGCAATGGAATGGTATC 180
    |||||||
QY 271 CTGCTATGGGTTACAGAGAAACTGAGAAGCATGACGAGCAAGTACAGGCTGACCTG 330
    |||||||
Db 181 CTGCTATGGGTTACAGAGAAACTGAGAAGCATGACGAGCAAGTACAGGCTGACCTG 240
    |||||||
QY 331 TCAGTGGCGACCTCCTCTTTGTCAATCAGCCTTCCTTGGGCAATTGATGCCGTGCA 390
    |||||||
Db 241 TCAGTGGCGACCTCCTCTTTGTCAATCAGCCTTCCTTGGGCAATTGATGCCGTGCA 300
    |||||||
QY 391 AACTGTACTTTGGGAACCTCTATGCAAGGAGTCCATGCTATCTACAGTCAACCTC 450
    |||||||
Db 301 AACTGTACTTTGGGAACCTCTATGCAAGGAGTCCATGCTATCTACAGTCAACCTC 360
    |||||||
QY 451 TACAGCAGTGTCTCTATCTCTGCGCTTATCATGCTGAGACCGCTACCTGGCATCTGCAC 510
    |||||||
Db 351 TACAGCAGTGTCTCTATCTCTGCGCTTATCATGCTGAGACCGCTACCTGGCATCTGCAC 420
    |||||||
QY 511 GCCACCAAGAGTAGAGGCCAAGAAAGCTGTGGCTGAAAAAGTGGTCTATGTTGGGCTC 570
    |||||||
Db 421 GCCACCAAGAGTAGAGGCCAAGAAAGCTGTGGCTGAAAAAGTGGTCTATGTTGGGCTC 480
    |||||||
QY 571 TGGATCCCTGCCCTCTCTGCTGACTATCTCCGACTTATCTTCCCAAGCTCAGTGAAGCA 630
    |||||||
Db 481 TGGATCCCTGCCCTCTCTGCTGACTATCTCCGACTTATCTTCCCAAGCTCAGTGAAGCA 540
    |||||||
QY 631 GATGACAGATATATCTGTGACCGCTTCTACCCCAATGACTTGTGGGTGTTGTGTTCCAG 690
    |||||||
Db 541 GATGACAGATATATCTGTGACCGCTTCTACCCCAATGACTTGTGGGTGTTGTGTTCCAG 600
    |||||||
QY 691 TTTCACACATCATGTTGGCTTATCTGCGCTGATTTGCTCATCTGCTGCTCTATGCTG 750
    |||||||
Db 601 TTTCACACATCATGTTGGCTTATCTGCGCTGATTTGCTCATCTGCTGCTCTATGCTG 660
    |||||||
QY 751 ATTATCATCTCCAAAGCTGTACACTCCAAAGGGGACAGAGGCAAGGCCCTCAAGACC 810
    |||||||
Db 661 ATTATCATCTCCAAAGCTGTACACTCCAAAGGGGACAGAGGCAAGGCCCTCAAGACC 720
    |||||||
QY 811 ACAGTCATCTCTATCTGCTTCTGCGCTGTTGGCTGCTTACTACATTTGGGATCAGC 870
    |||||||
Db 721 ACAGTCATCTCTATCTGCTTCTGCGCTGTTGGCTGCTTACTACATTTGGGATCAGC 780
    |||||||
QY 871 ATGCACTCTCTATCTGCTTCTGGAATCATCAAGAGGGGTGAGTTTGGAGACACGTG 930
    |||||||
Db 781 ATGCACTCTCTATCTGCTTCTGGAATCATCAAGAGGGGTGAGTTTGGAGACACGTG 840
    |||||||
QY 931 CACAAGTGAATTCATCAGCGAGGCCCTAGCTTCTTCCACTGTTGTCTGAAGCCCATC 990
    |||||||
Db 841 CACAAGTGAATTCATCAGCGAGGCCCTAGCTTCTTCCACTGTTGTCTGAAGCCCATC 900
    |||||||
QY 991 CTCTATGCTTCTGAGGCAAAATTTAAAGCTGCGCCAGCAGCAGCTACCTCTGTG 1050
    |||||||
Db 901 CTCTATGCTTCTGAGGCAAAATTTAAAGCTGCGCCAGCAGCAGCTACCTCTGTG 960
    |||||||
QY 1051 ACAGAGGGTCCAGGCTCAAGATCTCTCCAAAGAAAGGAGGTGAGACATTCATCTGT 1110
    |||||||
  
```

Db 961 AGCAGAGGTCACAGCCTCAAGATCCTCTCCAAAGAAAGCAGAGTGACATTCATCTGTT 1020

QY 1111 TCCACTGAGTCTGAGTCTTCAAGTTTCACTCCAGCTTAA 1149
 |||||||

Db 1021 TCCACTGAGTCTGAGTCTTCAAGTTTCACTCCAGCTTAA 1059

RESULT 15
 PTU89798
 LOCUS PTU89798 1087 bp mRNA PRI 23-FEB-1998
 DEFINITION Pan troglodytes CXK chemokine receptor 4 (CXCR4) mRNA, complete cds.
 ACCESSION U89798
 NID 92281446
 VERSION U89798.1 GI:2281446
 KEYWORDS
 SOURCE chimpanzee.
 ORGANISM Pan troglodytes
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pan.
 REFERENCE 1 (bases 1 to 1087)
 Pretet,J.L., Zerbib,A.C., Girard,M., Guillet,J.G., and Butor,C.
 TITLE Chimpanzee CXCR4 and CCR5 act as coreceptors for HIV type 1
 JOURNAL AIDS Res. Hum. Retroviruses 13 (18), 1583-1597 (1997)
 MEDLINE 98090115
 REFERENCE 2 (bases 1 to 1087)
 Pretet,J.-L., Zerbib,A., Girard,M., Guillet,J.-G., and Butor,C.
 TITLE Direct Submission
 AUTHORS Submitted (17-FEB-1997) U445, INSERM, ICGM, 22 rue Mechain, Paris 75014, France

FEATURES
 source location/Qualifiers
 1..1087
 /organism="Pan troglodytes"
 /db_xref="taxon:9598"
 /cell_type="PBL"
 1..1087
 /gene="CXCR4"
 29..1087
 /note="CXCR4"
 /gene="CXCR4"
 /note="CXCR4; a seven-transmembrane domain receptor"
 /codon_start=1
 /product="CXK chemokine receptor 4"
 /protein_id="AAC03718.1"
 /db_xref="GI:2281447"
 /db_xref="PIR:G2281447"
 /translation="MEGISITYTSDNTEEMSGSDYDSMKPECFREBNAMNKIFLPTI
 VSILFLTGIVGNGIVILVMGYOKKLRSMTEKDYRIHLHVSVDLLFVILIPWADAVANM
 YGNFLCKRAVHYITVNLVSSVLIILAFISLDRIALVAHATNSORPKLLAEVYVGV
 WIPALDLTPDITFANVSEADRDYICDRYPNDLVVAFQHIMVGLIPGIVILSC
 YCIILSKISHSGHOKRKALKTTVLILAFACWLPYVIGISIDSLILEIIRKGEF
 ENTVHKMISITLALAFHCLNPLIYAFILAFAGAKFTSAOHALTSVSGSLIKLSKGR
 GHSVSYSSTSESSPFS"
 BASE COUNT 250 a 311 c 239 g 287 t

ORIGIN

Query Match 56.7%; Score 985; DB 11; Length 1087;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1085; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 63 CACCGCATCTGGGGAACACAGCGGTATACATGAGGGGATCAGTATATACACTTCAGATAA 122
    |||||||
Db 1 CACCGCATCTGGGGAACACAGCGGTATACATGAGGGGATCAGTATATACACTTCAGATAA 60
    |||||||
QY 123 CTACACCGAGGAATGGGCTCAGGGAGCTATGACTCATGAAGAACCCGTTCGTTG 182
    |||||||
Db 61 CTACACCGAGGAATGGGCTCAGGGAGCTATGACTCATGAAGAACCCGTTCGTTG 120
    |||||||
QY 183 AGAAAAATGCTAATTTCAATAAATCTTCTGCCACCATCTACTCCATCATCTTTTAA 242
    |||||||
Db 121 AGAAAAATGCTAATTTCAATAAATCTTCTGCCACCATCTACTCCATCATCTTTTAA 180
    |||||||
QY 243 TGGCATTTGGGGCAATGGATTGGTCACTCCTGCTCATGGGTTACCAAGAAACTGAGAG 302
    |||||||
  
```

```

181 TGGCATGTGGGCAATGATGGTGCATCTGGTGCATGGGTACAGAAAGCAAGGAGAA 240
QY 303 CATGACGGAAGAGACAGGCTGCACCTGTAGTGGCCGACCTCTCTTTGTATCAGCT 362
Db 241 CATGACGGAAGAGACAGGCTGCACCTGTAGTGGCCGACCTCTCTTTGTATCAGCT 300
QY 363 TCCCTTCTGGGCAAGTGTATCCCGTGGCAAACTGGTACTTTGGAACTTCTTATGCAAGG 422
Db 301 TCCCTTCTGGGCAAGTGTATCCCGTGGCAAACTGGTACTTTGGAACTTCTTATGCAAGG 360
QY 423 AGTCATGTATCTACACAGCTCAACCTCTACAGAGAGTGTCTTCACTGGCTTCATCAG 482
Db 361 AGTCATGTATCTACACAGCTCAACCTCTACAGAGAGTGTCTTCACTGGCTTCATCAG 420
QY 483 TCTGGACCGGTACTGGCCATCGTCCACGCCACCAAGTCAGAGGCAAGAACTGTT 542
Db 421 TCTGGACCGGTACTGGCCATCGTCCACGCCACCAAGTCAGAGGCAAGAACTGTT 480
QY 543 GGCCTGAAAAGGTGTCTATGTGGCGCTGGATCCCTGCCCTCGTGTACTATTCGCGA 602
Db 481 GGCCTGAAAAGGTGTCTATGTGGCGCTGGATCCCTGCCCTCGTGTACTATTCGCGA 540
QY 603 CTTCATCTTGTCCACAGCTCAGTGAAGGAGATGACAGATATATCTGTACCGCTTACCC 662
Db 541 CTTCATCTTGTCCACAGCTCAGTGAAGGAGATGACAGATATATCTGTACCGCTTACCC 600
QY 663 CAATGACTTGTGGGTGTGTGTTCAGTTTCAGACATATATGTTGGCTTATTCCTGCC 722
Db 601 CAATGACTTGTGGGTGTGTGTTCAGTTTCAGACATATATGTTGGCTTATTCCTGCC 660
QY 723 TGGTATGTATCTCTGTCTGTCTATTCATATTCATCTCCAACTGTACACTCCAAAGG 782
Db 661 TGGTATGTATCTCTGTCTGTCTATTCATATTCATCTCCAACTGTACACTCCAAAGG 720
QY 783 CCACCAAGAGCGGAAGGCCCTCAAGACACAGTATCTCTATCTGTCTGTCTGTG 842
Db 721 CCACCAAGAGCGGAAGGCCCTCAAGACACAGTATCTCTATCTGTCTGTCTGTG 780
QY 843 TTGGCTGCTTACTATGATGGGATGGGATGAGATGAGTCTCTCTCTCTGTGAAATTCATCA 902
Db 781 TTGGCTGCTTACTATGATGGGATGGGATGAGATGAGTCTCTCTCTGTGAAATTCATCA 840
QY 903 GCAAGGGGTGAGTTTGAGAACACTGTGCACAAAGTGGATTCATCACCAGGCGCTAGC 962
Db 841 GCAAGGGGTGAGTTTGAGAACACTGTGCACAAAGTGGATTCATCACCAGGCGCTAGC 900
QY 963 TTTCTTCCACTGTGTGTCTGAACCCCATCTCTATGCTTTCTTCTGGAGCCAAATTTAAAC 1022
Db 901 TTTCTTCCACTGTGTGTCTGAACCCCATCTCTATGCTTTCTTCTGGAGCCAAATTTAAAC 960
QY 1023 CTCTGCCAGCAGCAGTCACTCTGTGTGAGCAGAGGGTCCAGGCTCAAGTCTCTCCAA 1082
Db 961 CTCTGCCAGCAGCAGTCACTCTGTGTGAGCAGAGGGTCCAGGCTCAAGTCTCTCCAA 1020
QY 1083 AGCAAGGAGGTGAGCATTCATCTTTTCCACTGAGTCTGAGTCTTCAAGTTTTCAGT 1142
Db 1021 AGCAAGGAGGTGAGCATTCATCTTTTCCACTGAGTCTGAGTCTTCAAGTTTTCAGT 1080
QY 1143 CAGCTAA 1149
Db 1081 CAGCTAA 1087

```

```

RESULT 16
MMU93311 1059 bp mRNA PRI 14-MAY-1997
LOCUS Macaca mulatta alpha chemokine receptor (CXCR4) mRNA, complete cds.
DEFINITION
ACCESSION U93311
VERSION 91934670
KEYWORDS U93311.1 GI:1934670
SOURCE rhesus monkey.

```

```

ORGANISM Macaca mulatta
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae;
Macaca.
REFERENCE 1 (bases 1 to 1059)
AUTHORS Edinger, A.L., Amedee, A., Miller, K., Doranz, B.J., Endres, M.,
Sharon, M., Samson, M., Lu, Z.-H., Clements, J.E., Murphy-Corb, M.,
Peifer, S.C., Parmentier, M., Broder, C.C. and Doms, R.W.
TITLE Differential utilization of CCR5 by macrophage and T cell tropic
simian immunodeficiency virus strains
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 94 (8), 4005-4010 (1997)
MEDLINE 97268687
REFERENCE 2 (bases 1 to 1059)
AUTHORS Miller, K.A., Hauer, D.A. and Clements, J.E.
TITLE Rhesus macaque CXCR4
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 1059)
AUTHORS Miller, K.A., Hauer, D.A. and Clements, J.E.
TITLE Direct Submission
JOURNAL Submitted (13-MAR-1997) Comparative Medicine, Johns Hopkins
University School of Medicine, 720 Rutland Ave., Ttaylor G-60,
Baltimore, MD 21205, USA
FEATURES
source
1..1059
Location/Qualifiers
/organism="Macaca mulatta"
/db_xref="taxon:9544"
/cell_type="PHA-biased normal PBMC"
1..1059
/gene="CXCR4"
1..1059
/note="human CXCR4 ortholog; fusin; LESTR; HUMSTR; LCR1; G
protein-coupled receptor"
/codon_start=1
/product="alpha chemokine receptor"
/protein_id="AAB54116.1"
/db_xref="PIR:G1934671"
/db_xref="GI:1934671"
translation="MEGISITSTNTEEMSGSDYSIKPECFRENAHFRILPTI
YSLFNGIVNGVILVMGYOEKLRSTEDKRIHLSTADLEFVITLPPRAVDVAM
KFGNFKAVAVITVNIYSSVLAFLSIRITAIYATVSPKRLAKVYVAG
WIPALLTIDPFIASVSEADRYICREFPNIDLVVFOHIMVGLILPGLVLS
CYILISLSHSGHOKRAKLTIVILAFPAWLPYIGISIDSLIELIKGCE
ENTVHKWISITELALFHCCLNPILVIAFLAKFYSQHALTVSRGSLKILSKR
GHSVSSTSSSNFHS"
BASE COUNT 241 a 301 c 236 g 281 t
ORIGIN
Query Match 19.6%; Score 341; DB 10; Length 1059;
Best Local Similarity 98.9%; Pred. No. 1e-173;
Matches 741; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 346 CTCTTTGTATCAGCTTCCCTTCTGGGCAAGTGTGATGCGGCAAACTGTACTTTGGG 405
Db 346 CTCTTTGTATCAGCTTCCCTTCTGGGCAAGTGTGATGCGGCAAACTGTACTTTGGG 315
QY 256 CTCTTTGTATCAGCTTCCCTTCTGGGCAAGTGTGATGCGGCAAACTGTACTTTGGG 315
Db 256 CTCTTTGTATCAGCTTCCCTTCTGGGCAAGTGTGATGCGGCAAACTGTACTTTGGG 315
QY 406 AACTTCCATGCAAGGAGTGCATGCTACAGAGCAACCTTACAGAGTCTCTC 465
Db 406 AACTTCCATGCAAGGAGTGCATGCTACAGAGCAACCTTACAGAGTCTCTC 375
QY 316 AACTTCCATGCAAGGAGTGCATGCTACAGAGCAACCTTACAGAGTCTCTC 375
Db 316 AACTTCCATGCAAGGAGTGCATGCTACAGAGCAACCTTACAGAGTCTCTC 375
QY 466 ATCTGAGCCTTCATGAGTGTGAGCGGTACAGGCGCATGCTCAAGCAGCAAGAGTCTAG 525
Db 466 ATCTGAGCCTTCATGAGTGTGAGCGGTACAGGCGCATGCTCAAGCAGCAAGAGTCTAG 435
QY 526 AGGCCAAGAGAGCTGTGGCTGAGAAAGTGTCTATGTTGGCGTGTGATCCCTGCCCTC 585
Db 526 AGGCCAAGAGAGCTGTGGCTGAGAAAGTGTCTATGTTGGCGTGTGATCCCTGCCCTC 495
QY 436 AAGCCAAAGAGAGCTGTGGCTGAGAAAGTGTCTATGTTGGAGTCTGATCCCTGCCCTC 495
Db 436 AAGCCAAAGAGAGCTGTGGCTGAGAAAGTGTCTATGTTGGAGTCTGATCCCTGCCCTC 495
QY 586 CTGTGACTATCCCGAGCTTCATCTTCCCAAGCGTGTGAGGAGAGATGACAGATATATC 645
Db 586 CTGTGACTATCCCGAGCTTCATCTTCCCAAGCGTGTGAGGAGAGATGACAGATATATC 555

```

```
QY 546 TGTACCGCTTCTACCCCAATGACTTGTGGTGTGTTCCAGTTTCACACATCATG 705
    |||||||
Db 556 TGTGACCGCTTCTACCCCAATGACTTGTGGTGTGTTCCAGTTTCACACATCATG 615
QY 706 GTTGGCTTATCTGCTGCTGATTTGTCATCTGTCTGCTATTGATTCATTCATCTCCAG 765
    |||||||
Db 616 GTTGGCTTATCTGCTGCTGATTTGTCATCTGTCTGCTATTGATTCATTCATCTCCAG 675
QY 766 CTCTCAGCTCCAGAGGCGCCAGAACCGGAGGCTCTCAAGACCAAGTCATCTCATC 825
    |||||||
Db 676 CTCTCAGCTCCAGAGGCGCCAGAACCGGAGGCTCTCAAGACCAAGTCATCTCATC 735
QY 826 CTGGCTTTCTTCCGCTGCTGCTGCTTACTAGATTGGATTCAGACGACTCCTTCATC 885
    |||||||
Db 736 CTGGCTTTCTTCCGCTGCTGCTGCTTACTAGATTGGATTCAGACGACTCCTTCATC 795
QY 886 CTCTCTGGAATATCATCAGCAGAGGCTGTGAGTTGAGAACACTGTGCACAAGTGGATTCC 945
    |||||||
Db 796 CTCTCTGGAATATCATCAGCAGAGGCTGTGAGTTGAGAACACTGTGCACAAGTGGATTCC 855
QY 946 ATCAGCAGAGGCTTCTTCCACTGTTCTGACCTGTTCTGACCCCATCTCTTCTTCTTCTT 1005
    |||||||
Db 856 ATACAGCAGAGGCTTCTTCCACTGTTCTGACCTGTTCTGACCCCATCTCTTCTTCTTCTT 915
QY 1006 GAGCCAAATTTAAACCTCTGCCCAGACGACGACTCCTCTGTGAGCAGAGGCTCCAGC 1065
    |||||||
Db 916 GAGCCAAATTTAAACCTCTGCCCAGACGACGACTCCTCTGTGAGCAGAGGCTCCAGC 975
QY 1066 CTCAGATCTCTCTCCCAAGGAAAGCGAGG 1094
    |||||||
Db 976 CTCAGATCTCTCTCCCAAGGAAAGCGAGG 1004
```

```
RESULT 17
AF001928 1087 bp mRNA PRI 10-MAR-1998
LOCUS Macaca mulatta alpha-chemokine receptor 4 (CXCR4) mRNA, complete
DEFINITION
ACCESSION AF001928
NID 92911293
VERSION AF001928.1 GI:2911293
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
SOURCE
```

```
1.1087
/organism="Macaca mulatta"
/strain="chinese"
/db_xref="taxon:9544"
/cell_type="peripheral blood lymphocyte"
1..1087
/gene="CXCR4"
29..1087
/feature="CXCR4"
/note="seven trans-membrane domain receptor"
/codon_start=1
/product="alpha-chemokine receptor 4"
/protein_id="AAC39641.1"
/db_xref="PID:92911294"
```

```
/db_xref="GI:2911294"
/translation="MEGISTYSDNTEPMGSDYDSEKPECFREBNAHFNPLEPTI
YSLIFLTGIVGVLIVMGYOKRLSMIDKRLHLISVADLLFVITLLPMAVDVAMW
YEGNEFLKRAVHVIVYVNLISVLIASFSLDYLIVAHVINSOKPKLLAEVYVGV
WIPALDLITPDFIFASVSEADRYICDRFVNDLVWVFOFQIMVGLILPGVILSC
YCIITSLKSHSGHOKRAKTTVTLIIAFAFCQWLPYGISIDSLILEITIKOGEF
ENLVHWSITETALAFHCCCLNPILVAFILGAIFKISQALHISVNSGSLKILSKGR
GGHSYSTSESSSRHSS"
```

```
BASE COUNT 246 a 310 c 245 g 286 t
ORIGIN
Query Match 19.6% Score 341 DB 11: Length 1087;
Best Local Similarity 98.9% Pred. No. 1e-173;
Matches 741; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 346 CTCTTTGTATCAGCGCTTCCCTTCTGCGCAGTTGATGCCGTGCAAACTGGTACTTTGGG 405
    |||||||
Db 284 CTCTTTGTATCAGCGCTTCCCTTCTGCGCAGTTGATGCCGTGCAAACTGGTACTTTGGG 343
QY 406 AACTTCTATGCAAGGAGTCCATGTATCTACACAGTCACTCTACACAGAGTCTCTC 465
    |||||||
Db 344 AACTTCTATGCAAGGAGTCCATGTATCTACACAGTCACTCTACACAGAGTCTCTC 403
QY 466 ATCTGAGCTTCAATCAGTGTGAGCGCTACCTGCGCATGCTCCAGCGCCACCAAGTCCAG 525
    |||||||
Db 404 ATCTGAGCTTCAATCAGTGTGAGCGCTACCTGCGCATGCTCCAGCGCCACCAAGTCCAG 463
QY 526 AGGCCAAGGAAGCTGTGGCTGAAAGGTGCTATGTTGGCGTCTGGATCCTGCGCTC 585
    |||||||
Db 464 AAGCCAAAGGAAGCTGTGGCTGAAAGGTGCTATGTTGGAGTCTGGATCCTGCGCTC 523
QY 586 CTGCTGACTATTTCCGCACTTATCTTGGCAAGTGTAGTGGCAATATGACATATATC 645
    |||||||
Db 524 CTGCTGACTATTTCCGCACTTATCTTGGCAAGTGTAGTGGCAATATGACATATATC 583
QY 646 TGTGACCGCTTCTTACCCCAATGACTTGTGGGTGCTGTGTCCAGTTTCAGCATATCATG 705
    |||||||
Db 584 TGTGACCGCTTCTTACCCCAATGACTTGTGGGTGCTGTGTCCAGTTTCAGCATATCATG 643
QY 706 GTTGGCTTATCTGCTGCTGATTTGTCATCTGTCTGCTGATTCATTCATCTCCAG 765
    |||||||
Db 644 GTTGGCTTATCTGCTGCTGATTTGTCATCTGTCTGCTGATTCATTCATCTCCAG 703
QY 766 CTCTCAGCTCCAGAGGCGCCAGAACCGGCTCAAGACCAAGTCACTCATCTCATC 825
    |||||||
Db 704 CTCTCAGCTCCAGAGGCGCCAGAACCGGCTCAAGACCAAGTCACTCATCTCATC 763
QY 826 CTGGCTTTCTTCCGCTGCTGCTGCTTACTAGATTGGATTCAGTCACTCTTCATC 885
    |||||||
Db 764 CTGGCTTTCTTCCGCTGCTGCTGCTTACTAGATTGGATTCAGTCACTCTTCATC 823
QY 886 CTCTCTGGAATATCATCAGCAGAGGCTGTGAGTTGAGAACACTGTGCACAAGTGGATTCC 945
    |||||||
Db 824 CTCTCTGGAATATCATCAGCAGAGGCTGTGAGTTGAGAACACTGTGCACAAGTGGATTCC 883
QY 946 ATCAGCAGAGGCTTCTTCCACTGTGTGTCGACCCCATCTCTATGTTCTTCTT 1005
    |||||||
Db 884 ATCAGCAGAGGCTTCTTCCACTGTGTGTCGACCCCATCTCTATGTTCTTCTT 943
QY 1006 GAGCCAAATTTAAACCTCTGCCCAGACGACGACTCCTGTGAGCAGAGGCTCCAGC 1065
    |||||||
Db 944 GAGCCAAATTTAAACCTCTGCCCAGACGACGACTCCTGTGAGCAGAGGCTCCAGC 1003
QY 1066 CTCAGATCTCTCTCCCAAGGAAAGCGAGG 1094
    |||||||
Db 1004 CTCAGATCTCTCTCCCAAGGAAAGCGAGG 1032
```

```
RESULT 18
D86579 1078 bp mRNA PRI 07-FEB-1999
LOCUS Macaca fascicularis mRNA for fusin (LESTR), complete cds.
DEFINITION
```

ACCESSION D86579
 MID g1468948
 VERSION D86579.1 GI:1468948
 KEYWORDS fusin (LESTR).
 SOURCE Macaca fascicularis cDNA to mRNA.
 ORGANISM Macaca fascicularis
 Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
 Primates; Catarrhini; Cercopithecoidea; Cercopithecinae; Macaca.
 REFERENCE 1 (bases 1 to 1078)
 AUTHORS Tatsumi, M.
 TITLE Direct Submission
 JOURNAL Submitted (18-JUL-1996) to the DDBJ/EMBL/GenBank databases. Masashi
 Tatsumi, National Institute of Health, Department of Veterinary
 Science, Toyama-23-1, Shinjuku-Ku, Tokyo 162, Japan
 (E-mail: tatsu@nih.go.jp, Tel: 81-3-5285-1111, Fax: 81-3-5285-1179)
 2 (sites)
 REFERENCE 2
 AUTHORS Monkey CD4 and fusin are not species barrier for HIV-1 replication
 TITLE Unpublished (1996)
 JOURNAL
 FEATURES
 source
 1. 1078
 Location/Qualifiers
 /organism="Macaca fascicularis"
 /db_xref="taxon:9541"
 11. 1069
 /codon_start=1
 /product="Fusin (LESTR)"
 /protein_id="BAAL3126.1"
 /db_xref="PID:d1013813"
 /db_xref="PID:g1468949"
 /db_xref="GI:1468948"
 /translation="MEGISIYTSNDNTEEMSGSDYSIKEPCRENAHNRILFPTI
 YSIFLGIIVNGVILVMGYOKRLSMIDRYLHLSVDLLFVILFPAVDVAVNM
 YFENFLCAVAVIYVNLVSSVILAFISDRYLAVHATNSRPRKLAEKVYVGV
 WIPALLITIDPFIASVSEADRYICDRFPNDLVWVVFQFQIHNGLLIPGVILSC
 YCIILSLSHSGHOKRAKLTIVILIAFAWMLPYVIGISIDSLILEIKOGCEP
 ENTVHKWISITELAFHCCINPLILVAFGAKRTISAQHALTSVSGSSILKLSKGR
 GGHSSVTESESSPFS"
 BASE COUNT 244 a 304 c 243 g 287 t
 ORIGIN

Query Match 19.3% Score 336; DB 9; Length 1078;
 Best Local Similarity 98.9%; Pred. No. 5.2e-171;
 Matches 736; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 351 TGTGATACGCTTCCTCTGGGCGAGTTGATGCCGTGCAAACTGTACTTTGGAACTT 410
 |||||||
 DB 271 TGTGATACGCTTCCTCTGGGCGAGTTGATGCCGTGCAAACTGTACTTTGGAACTT 330
 |||||||
 QY 411 CCTATGCAAGGAGTCATGTCATACAGTCAACCTCTACAGCAGTGTCTCATCT 470
 |||||||
 DB 331 CCTGTCAGGAGTCATGTCATACAGTCAACCTCTACAGCAGTGTCTCATCT 390
 |||||||
 QY 471 GGCCTTCATGCTGAGCGCTACCTGGCCATGTCACAGCCACCAAGTACAGAGCC 530
 |||||||
 DB 391 GGCCTTCATGCTGAGCGCTACCTGGCCATGTCACAGCCACCAAGTACAGAGCC 450
 |||||||
 QY 531 AAGGAGCTTTGGCTGAAAAGTGTATGTTGGGCTGTGATCCCTGCTGCT 590
 |||||||
 DB 451 AAGGAGCTTTGGCTGAAAAGTGTATGTTGGGCTGTGATCCCTGCTGCT 510
 |||||||
 QY 591 GACTATTCGCGAGCTTCATCTTTCGCAACGTCATGAGGCGAGATGACAGATATATCTGTGA 650
 |||||||
 DB 511 GACTATTCGCGAGCTTCATCTTTCGCAACGTCATGAGGCGAGATGACAGATATATCTGTGA 570
 |||||||
 QY 651 CCAGCTTCACCCCAATGACTTGTGGTGTGTGTCCAGTTTCAGACATCATGTTGG 710
 |||||||
 DB 571 CCAGCTTCACCCCAATGACTTGTGGTGTGTGTCCAGTTTCAGACATCATGTTGG 630
 |||||||
 QY 711 CCTATTCGCTGCTGATTTGTCTATCTGTCTGCTATATGCAATATCATCTCAAGCTATC 770
 |||||||
 DB 631 CCTATTCGCTGCTGATTTGTCTATCTGTCTGCTATATGCAATATCATCTCAAGCTATC 690
 |||||||

QY 771 AACTCCAAAGGCGCCACGACGAAGGCGCCCTCAAGACCATCATCTCTAGC 830
 |||||||
 DB 691 AACTCCAAAGGCGCCACGACGAAGGCGCCCTCAAGACCATCATCTCTAGC 750
 |||||||
 QY 831 TTCTTCGCTGTGGCTGCTTACTACATTTGGATCAGATGATCCCTCATCTCT 890
 |||||||
 DB 751 TTCTTCGCTGTGGCTGCTTACTACATTTGGATCAGATGATCCCTCATCTCT 810
 |||||||
 QY 891 GGAATATCAAGCAAGGCTGTAGTTTGAGAACACTGTGCACAACTGTATTCATC 950
 |||||||
 DB 811 GGAATATCAAGCAAGGCTGTAGTTTGAGAACACTGTGCACAACTGTATTCATC 870
 |||||||
 QY 951 CGAGGCGCTGCTTCTTCACAGTGTGCAGCCCATCTCTATCTCTCTTCTTGGAG 1010
 |||||||
 DB 871 CGAGGCGCTGCTTCTTCACAGTGTGTGAGCCCATCTCTATGCTTCTTGGAG 930
 |||||||
 QY 1011 CAATTTAAACCTGTCCAGCAGCAGCTCACTCTGTGAGCAGAGGCTCCAGCTCA 1070
 |||||||
 DB 931 CAATTTAAACCTGTCCAGCAGCAGCTCACTCTGTGAGCAGAGGCTCCAGCTCA 990
 |||||||
 QY 1071 GATCCTCTCCAAAGGAAGCGAGG 1094
 |||||||
 DB 991 GATCCTCTCCAAAGGAAGCGAGG 1014
 |||||||

RESULT 19
 LOCUS AF019378
 DEFINITION Cercopithecus aethiops G-protein coupled receptor (CXCR4) mRNA,
 complete cds.
 ACCESSION AF019378
 MID AF019378
 VERSION AF019378.1 GI:4102991
 KEYWORDS
 SOURCE African green monkey.
 ORGANISM Chlorocebus aethiops
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae;
 Chlorocebus.
 REFERENCE 1 (bases 1 to 1126)
 AUTHORS Holtkamp, N., Baier, M., and Werner, A.
 TITLE CXCR4 from African green monkey
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1126)
 AUTHORS Holtkamp, N., Baier, M., and Werner, A.
 TITLE Direct Submission
 JOURNAL Submitted (14-AUG-1997) 6/3, Paul-Ehrlich-Institut,
 Paul-Ehrlich-Str. 51-59, Langen 63225, Germany
 FEATURES
 source
 1. 1126
 Location/Qualifiers
 /organism="Chlorocebus aethiops"
 /db_xref="taxon:9534"
 1. 1126
 /gene="CXCR4"
 68. 1126
 /gene="CXCR4"
 /codon_start=1
 /product="G-protein coupled receptor"
 /protein_id="AA01638.1"
 /db_xref="PID:g4102992"
 /db_xref="GI:4102991"
 /translation="MEGISIYTSNDNTEEMSGSDYSIKEPCRENAHNRILFPTI
 YSIFLGIIVNGVILVMGYOKRLSMIDRYLHLSVDLLFVILFPAVDVAVNM
 YFENFLCAVAVIYVNLVSSVILAFISDRYLAVHATNSRPRKLAEKVYVGV
 WIPALLITIDPFIASVSEADRYICDRFPNDLVWVVFQFQIHNGLLIPGVILSC
 YCIILSLSHSGHOKRAKLTIVILIAFAWMLPYVIGISIDSLILEIKOGCEP
 ENTVHKWISITELAFHCCINPLILVAFGAKRTISAQHALTSVSGSSILKLSKGR
 GGHSSVTESESSPFS"
 BASE COUNT 254 a 318 c 260 g 294 t
 ORIGIN

Query Match 18.0% Score 313; DB 11; Length 1126;

Best Local Similarity 98.8%; Pred. No. 1.5e-158;
Matches 813; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```
QY 272 TGGTCATGGGTACCAAGAAAGTGAAGCATGACGCAAGTACAGGCTCACCTGT 331
    |||||||
Db 249 TGGTCATGGGTACCAAGAAAGTGAAGCATGACGCAAGTACAGGCTCACCTGT 308
QY 332 CAGTGGCCAGCTCTCTTGTTCATCAGCCTTCCCTTGGGGAGTGTGATGCCGTGGCAA 391
    |||||||
Db 309 CAGTGGCTACCTCTCTTGTTCATCAGCCTTCCCTTGGGGAGTGTGATGCCGTGGCAA 368
QY 392 ACTGGTACTTGGGAATCTCTATGCAAGGACGATCATCTATACAGAGTCAACCTCT 451
    |||||||
Db 369 ACTGGTACTTGGGAATCTCTATGCAAGGACGATCATCTATACAGAGTCAACCTCT 428
QY 452 ACAGAGTGTCTCATCTCTGGGCTTATCAGTGTGACGCGTACCTGGGCTATGCTCCAG 511
    |||||||
Db 429 ACAGAGTGTCTCATCTCTGGGCTTATCAGTGTGACGCGTACCTGGGCTATGCTCCAG 488
QY 512 CCACCAACAGTCAGAGGCGCAAGAGCTTGTGGCTGAAAGTGTCATATGTGGCTCT 571
    |||||||
Db 489 CCACCAACAGTCAGAGGCGCAAGAGCTTGTGGCTGAAAGTGTCATATGTGGCTCT 548
QY 572 GGATCCCTGCTCTCTGCTGACTATTCGCGACTTCATCTTTGCCAAGCTCAGTGAGGCGAG 631
    |||||||
Db 549 GGATCCCTGCTCTCTGCTGACTATTCGCGACTTCATCTTTGCCAAGCTCAGTGAGGCGAG 608
QY 632 ATGACAGATATATCTGTGACGCTTCTACCCCAATGACTTGGGGTGTGTTCAGT 691
    |||||||
Db 609 ATGACAGATATATCTGTGACGCTTCTACCCCAATGACTTGGGGTGTGTTCAGT 668
QY 692 TTCACACATCATGTTGGCTTTATCTGCTGCTGATTTGTCTATCTGCTGCTATTTGCA 751
    |||||||
Db 669 TTCACACATCATGTTGGCTTTATCTGCTGCTGATTTGTCTATCTGCTGCTATTTGCA 728
QY 752 TTATCATCTCCAAAGCTGTACACTCCAAAGGCGCACGAAAGGCGCTTCAAGACCA 811
    |||||||
Db 729 TTATCATCTCCAAAGCTGTACACTCCAAAGGCGCACGAAAGGCGCTTCAAGACCA 788
QY 812 CAGTATCTCTCATCTCTGCTTCTGCGCTGTGGCTGCTTACTACTATGGGATAGCA 871
    |||||||
Db 789 CAGTATCTCTCATCTCTGCTTCTGCGCTGTGGCTGCTTACTACTATGGGATAGCA 848
QY 872 TCGACTCTCTCATCTCTCTGGAATCATCAAGCAAGGCTGTGAGTTTGAACACATGTC 931
    |||||||
Db 849 TCGACTCTCTCATCTCTCTGGAATCATCAAGCAAGGCTGTGAGTTTGAACACATGTC 908
QY 932 ACAAGTGGATTTCCATCAACGAGGCGCTAGCTTTCTTCACAGTGTGTGAACCCCATCC 991
    |||||||
Db 909 ACAAGTGGATTTCCATCAACGAGGCGCTAGCTTTCTTCACAGTGTGTGAACCCCATCC 968
QY 992 TCTATGCTTTCTTGGAGCAATTTAAACCTTGCCAGAGGAGCACTCAGCTGTGTA 1051
    |||||||
Db 969 TCTATGCTTTCTTGGAGCAATTTAAACCTTGCCAGAGGAGCACTCAGCTGTGTA 1028
QY 1052 GCAGAGGCTCAGGCTCAAGATCTCTCTCCAAAGGAAAGCGAGG 1094
    |||||||
Db 1029 GCAGAGGCTCAGGCTCAAGATCTCTCTCCAAAGGAAAGCGAGG 1071
```

RESULT 20
AF031089 1059 bp mRNA PRI 14-OCT-1998
LOCUS AF031089 Papio hamadryas anubis CXC chemokine receptor 4 (CXC4) mRNA,
DEFINITION complete cds.
ACCESSION AF031089
MID 2625093
VERSION AF031089.1 GI:2625093
KEYWORDS
SOURCE olive baboon.
ORGANISM Papio hamadryas anubis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae;

```
REFERENCE 1 (bases 1 to 1059)  
AUTHORS Benton, P.A., Lee, D.R., and Kennedy, R.C.  
TITLE Sequence comparisons of non-human primate HIV-1 coreceptor  
homologues  
JOURNAL Mol. Immunol. 35 (2), 95-101 (1998)  
MEDLINE 98346785  
REFERENCE 2 (bases 1 to 1059)  
AUTHORS Benton, P.A. and Kennedy, R.C.  
TITLE Direct Submission  
JOURNAL Submitted (22-Oct-1997) Microbiology and Immunology, University of  
Oklahoma Health Sciences Center, 800 Research Parkway, Rm 458,  
Oklahoma City, OK 73104, USA  
FEATURES  
source  
1..1059  
/organism="Papio hamadryas anubis"  
/sub_species="anubis"  
/db_xref="taxon:9555"  
/cell_type="peripheral blood mononuclear cells"  
1..1059  
/gene="CXC4"  
1..1059  
/gene="CXC4"  
/codon_start=1  
/product="CXC chemokine receptor 4"  
/protein_id="AAC63831.1"  
/db_xref="pid:92625094"  
/db_xref="gi:2625094"  
1..1059  
/translation="MEGISITSDNYTEEMSGSDYSIKEDCFRENAHFRIFLPTI  
YSILFLNGIVNGILVLMGYOKRLRSMTDYLRLHLYADLLFVTLTFPMADAVANM  
YEFNLEKCAHVIYTVNLSSVLIATFSLDRYLAVATNSQRPRLAKRVYVGV  
WIPALLTIPIEFASVEADRDYICDFEINDLVVYFOHIVGILGIVLSC  
YCIISKISHSGHOKRKALRTVLLIATFACWLPYIGISIFSLIETIKOCCEP  
ENTYHKMSTIDALAFHCCNPLIATLQKFKRTSAQHALTVSRGSLKILSKGR  
GSHSVSTSESSSPHSS"  
BASE COUNT 237 a 301 c 238 g 283 t  
ORIGIN  
Query Match 18.0%; Score 313; DB 11; Length 1059;  
Best Local Similarity 98.8%; Pred. No. 1.5e-158;  
Matches 813; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
QY 272 TGGTCATGGGTACCAAGAAAGTGAAGCATGACGCAAGTACAGGCTCACCTGT 331  
    |||||||  
Db 182 TGGTCATGGGTACCAAGAAAGTGAAGCATGACGCAAGTACAGGCTCACCTGT 241  
QY 332 CAGTGGCCAGCTCTCTTGTTCATCAGCCTTCCCTTGGGGAGTGTGATGCCGTGGCAA 391  
    |||||||  
Db 242 CAGTGGCTACCTCTCTTGTTCATCAGCCTTCCCTTGGGGAGTGTGATGCCGTGGCAA 301  
QY 392 ACTGGTACTTGGGAATCTCTATGCAAGGACGATCATCTATACAGAGTCAACCTCT 451  
    |||||||  
Db 302 ACTGGTACTTGGGAATCTCTATGCAAGGACGATCATCTATACAGAGTCAACCTCT 361  
QY 452 ACAGAGTGTCTCATCTCTGGGCTTATCAGTGTGACGCGTACCTGGGCTATGCTCCAG 511  
    |||||||  
Db 362 ACAGAGTGTCTCATCTCTGGGCTTATCAGTGTGACGCGTACCTGGGCTATGCTCCAG 421  
QY 512 CCACCAACAGTCAGAGGCGCAAGAGCTTGTGGCTGAAAGTGTCATATGTGGCTCT 571  
    |||||||  
Db 422 CCACCAACAGTCAGAGGCGCAAGAGCTTGTGGCTGAAAGTGTCATATGTGGCTCT 481  
QY 572 GGATCCCTGCTCTCTGCTGACTATTCGCGACTTCATCTTTGCCAAGCTCAGTGAGGCGAG 631  
    |||||||  
Db 482 GGATCCCTGCTCTCTGCTGACTATTCGCGACTTCATCTTTGCCAAGCTCAGTGAGGCGAG 541  
QY 632 ATGACAGATATATCTGTGACGCTTCTACCCCAATGACTTGGGGTGTGTGCTCCAGT 691  
    |||||||  
Db 542 ATGACAGATATATCTGTGACGCTTCTACCCCAATGACTTGGGGTGTGTGCTCCAGT 601  
QY 692 TTCACACATATGTTGGCTTTATCTGCTGCTGATATCTGCTGCTATTTGCA 751  
    |||||||
```

LOCUS	DEFINITION	ACCESSION	NID	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES	CDS
602	TTAGCAGCATATGTTGGCTTATCTCTGCTGGTATGATCCATCCGTCGCTGCTATTTGCA	661											
QY	752 TTATCATCTCTCCAGCTGTCTACACTCCAGGCGCCACCGAAGCGCAAGCCCTTCAAGACA	811											
Db	662 TTATCATCTCTCCAGCTGTCTACACTCCAGGCGCCACCGAAGCGCGCCCTTCAAGACA	721											
QY	812 CAGTCATCTCTATCTGGCTTTCTTGCCCTGTGGCGCTTACTACTATTTGGATGACA	871											
Db	722 CGGTCATCTCTATCTGGCTTTCTTGCCCTGTGGCGCTTACTACTATTTGGATGACA	761											
QY	872 TCGACTCTCTATCTCTGCTGGAAATCATCATCAAGCAAGGGTGTGATTTGAGAACACTGTGC	931											
Db	782 TCGACTCTCTATCTCTGCTGGAAATCATCATCAAGCAAGGGTGTGATTTGAGAACACTGTGC	841											
QY	932 ACAAGTGGATTTCCATCAGCAGGCGCCCTAGCTTTCTTCCACTGTGTTCTGAACCCATCC	991											
Db	842 ACAAGTGGATTTCCATCAGCAGGCGCCCTAGCTTTCTTCCACTGTGTTCTGAACCCATCC	901											
QY	992 TCTATGCTTTCTCTTGAGAGCCAAATTTAAACCTCTGCCGACGACGACACTACCTCTGTGA	1051											
Db	902 TCTATGCTTTCTCTTGAGAGCCAAATTTAAACCTCTGCCGACGACGACACTACCTCTGTGA	961											
QY	1052 GCACAGGGTCCAGGCTCAGATCCTCTCCAAAGGAAAGCGAG	1094											
Db	962 GCACAGGGTCCAGGCTCAGATCCTCTCCAAAGGAAAGCGAG	1004											
RESULT 21													
AF051906	AF051906 1059 bp mRNA	PRI	03-JUL-1998										
LOCUS	Cercopithecus torquatus atys PUG chemokine receptor CXCR4 mRNA,												
DEFINITION	complete cds.												
ACCESSION	AF051906												
NID	93135303												
VERSION	AF051906.1												
KEYWORDS	GI:3135303												
SOURCE	Sooty mangabey.												
ORGANISM	Cercopithecus torquatus atys												
REFERENCE	1 (bases 1 to 1059)												
AUTHORS	Chen,Z., Gettine,A., Ho,D.D. and Marx,P.A.												
TITLE	Primary SIVm isolates use the CCR5 coreceptor from sooty mangabeys												
JOURNAL	naturally infected in west Africa: a comparison of coreceptor usage												
FEATURES	of Primary SIVm, HIV-2, and SIVmac												
SOURCE	Virolology 246 (1), 113-124 (1998)												
FEATURES	2 (bases 1 to 1059)												
SOURCE	Chen,Z.												
FEATURES	Direct Submission												
SOURCE	Submitted (03-MAR-1998) Aaron Diamond AIDS Research Center, 455 1st												
FEATURES	Ave. 7th Floor, NYC, NY 10016, USA												
SOURCE	Location/Qualifiers												
FEATURES	1..1059												
SOURCE	/organism="Cercopithecus torquatus atys"												
FEATURES	/db_xref="taxon:9531"												
SOURCE	/cell_type="lymphocytes"												
FEATURES	/note="Sooty mangabey PUG; from uninfected animal"			</									

ORIGIN

Query Match	13.8%	Score 239;	DB 11;	Length 1059;
Best Local Similarity	100.0%	Pred. NO. 1.6e-118;		
Matches 239;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	856	TACATTGGGATCAGCATCGACTCCTCTTCATCTCTCTGGAATATCATACCAAGGGGTGAG	915
Db	766	TACATTGGGATCAGCATCGACTCCTCTTCATCTCTCTGGAATATCATACCAAGGGGTGAG	825
Qy	916	TTTGAAACACTGTGGCAAGTGGATTTTCATACCGAGGCCCTAGCTTCTTCACACT	975
Db	826	TTTGGAACACTGTGCACAGTGGATTTTCATACCGAGGCCCTAGCTTCTTCACACT	885
Qy	976	TGTCTGAACCCCATCTCTATGCTTTCTCTTGAGCCAAATTTAAACCTCTGCCAGCAC	1033
Db	886	TGTCTGAACCCCATCTCTATGCTTTCTCTTGAGCCAAATTTAAACCTCTGCCAGCAC	945
Qy	1036	GCATCTACCTCTGTGACACAAAGGGTCAGCTCTAATATCTCTCCAAAGGAAGCGAG	1094
Db	946	GCATCTACCTCTGTGACACAAAGGGTCAGCTCTAATATCTCTCCAAAGGAAGCGAG	1004

RESULT	22
MNCU73740	
LOCUS	MNCU73740 1059 bp mRNA PRI
DEFINITION	Macaca mulatta CXCR4 mRNA, complete cds.
ACCESSION	U73740
NID	G1771982
VERSION	U73740.1 GI:1771982
KEYWORDS	
SOURCE	rhesus monkey.
ORGANISM	Macaca mulatta

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES

1 (bases 1 to 1059)
Chen,Z., Zhou,P., Ho,D.D., Iandau,N.R. and Marx,P.A.
Genetically divergent strains of Simian immunodeficiency virus use
CCR5 as a coreceptor for entry
J. Virol. 71 (4), 2705-2714 (1997)
9723934
2 (bases 1 to 1059)
Chen,Z. and Marx,P.A.
Direct Submission
Submitted (08-OCT-1996) AIDS Animal Model, Aaron Diamond AIDS
Research Center, 455 1st Ave., 7th Floor, New York, NY 10016, USA
Location/Qualifiers

```

Source
1.103991="Macaena maculata"
/organism="indian macaque"
/strain="tacon 5514"
/di_xref="tacon 5514"
/cell_type="PBMC"
1.1059
/gene="CXCR4"
1.1059
CDS

```

BASE COUNT	241 a	300 c	236 g	282 t
ORIGIN				

BASE COUNT	241 a	299 c	234 g	285 t
1	1	1	1	1
2	1	1	1	1
3	1	1	1	1
4	1	1	1	1
5	1	1	1	1
6	1	1	1	1
7	1	1	1	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12	1	1	1	1
13	1	1	1	1
14	1	1	1	1
15	1	1	1	1
16	1	1	1	1
17	1	1	1	1
18	1	1	1	1
19	1	1	1	1
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	1	1	1	1
27	1	1	1	1
28	1	1	1	1
29	1	1	1	1
30	1	1	1	1
31	1	1	1	1
32	1	1	1	1
33	1	1	1	1
34	1	1	1	1
35	1	1	1	1
36	1	1	1	1
37	1	1	1	1
38	1	1	1	1
39	1	1	1	1
40	1	1	1	1
41	1	1	1	1
42	1	1	1	1
43	1	1	1	1
44	1	1	1	1
45	1	1	1	1
46	1	1	1	1
47	1	1	1	1
48	1	1	1	1
49	1	1	1	1
50	1	1	1	1
51	1	1	1	1
52	1	1	1	1
53	1	1	1	1
54	1	1	1	1
55	1	1	1	1
56	1	1	1	1
57	1	1	1	1
58	1	1	1	1
59	1	1	1	1
60	1	1	1	1
61	1	1	1	1
62	1	1	1	1
63	1	1	1	1
64	1	1	1	1
65	1	1	1	1
66	1	1	1	1
67	1	1	1	1
68	1	1	1	1
69	1	1	1	1
70	1	1	1	1
71	1	1	1	1
72	1	1	1	1
73	1	1	1	1
74	1	1	1	1
75	1	1	1	1
76	1	1	1	1
77	1	1	1	1
78	1	1	1	1
79	1	1	1	1
80	1	1	1	1
81	1	1	1	1
82	1	1	1	1
83	1	1	1	1
84	1	1	1	1
85	1	1	1	1
86	1	1	1	1
87	1	1	1	1
88	1	1	1	1
89	1	1	1	1
90	1	1	1	1
91	1	1	1	1
92	1	1	1	1
93	1	1	1	1
94	1	1	1	1

241 a	300 c	236 g	282 t
-------	-------	-------	-------

Query Match 13.8%; Score 239; DB 11; Length 1059;
Best Local Similarity 100.0%; Pred. No. 1.6e-118;
Matches 239: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 856 TACATTGGATAGCATGACATGCTTTCATCTCTGGAATATCATCAGCAGGGTGTAG 915
DB 756 TACATTGGATAGCATGACATGCTTTCATCTCTGGAATATCATCAGCAGGGTGTAG 825
OY 916 TTGAGAACCTGTGCACAGTGTGATTCATCAGCAGGCCCTTCTTCCACTGT 975
DB 826 TTGAGAACCTGTGCACAGTGTGATTCATCAGCAGGCCCTTCTTCCACTGT 885
OY 976 TGTCTGAACCCCATCTCTATGCTTCTTCTTGGAGCCAAATTTAAACCTCTGCCAGCAC 1035
DB 886 TGTCTGAACCCCATCTCTATGCTTCTTCTTGGAGCCAAATTTAAACCTCTGCCAGCAC 945
OY 1036 GCATCTACCTCTGTGAGCAGAGGGTTCACGCTCAGATCTCTCCAAAGAAAGCAGG 1094
DB 946 GCATCTACCTCTGTGAGCAGAGGGTTCACGCTCAGATCTCTCCAAAGAAAGCAGG 1004

RESULT 23

LOCUS AB015943 1068 bp mRNA PRI 06-FEB-1999
DEFINITION Chlorocebus aethiops mRNA for CXCR4 receptor, complete cds.
ACCESSION AB015943
NID 93294341
VERSION AB015943.1 GI:3294341
KEYWORDS CXCR4 receptor.
SOURCE Chlorocebus aethiops cDNA to mRNA.
ORGANISM Chlorocebus aethiops
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Cercopithecoidea; Cercopithecinae;
Chlorocebus.

REFERENCE 1 (bases 1 to 1068)
AUTHORS Murayama, Y., Matsunaga, S. and Inoue-Murayama, M.
TITLE Direct Submission
JOURNAL Submitted (01-JUL-1998) to the DDBJ/EMBL/GenBank databases. Yuichi Murayama, National Institute of Animal Health, Food Safety Research Laboratory, Kannondai 3-1-1, Tsukuba, Ibaraki 305-0856, Japan (E-mail:mspock@pc.mhn.or.jp, Tel:+81-298-38-7840, Fax:+81-298-38-7825)

REFERENCE 2 (bases 1 to 1068)
AUTHORS Murayama, Y., Matsunaga, S. and Inoue-Murayama, M.
TITLE cDNA sequence of African green monkey CXCR-4 chemokine receptor gene

JOURNAL Published Only in Database (1998) In press
FEATURES
SOURCE Location/Qualifiers
1..1068
/organism="Chlorocebus aethiops"
/db_xref="taxon:9534"
/note="Cercopithecus aethiops"
1..1059
/codon_start=1
/product="CXCR4 receptor"
/protein_id="BAA31327.1"
/db_xref="PID:G1032261"
/db_xref="PID:G13294342"
/db_xref="GI:3294342"
/translation="MEGISITSDNTTEEMSGDYDSIKKEPCFRENAHNRIFLPTI
YSITFLTGIVNGLVILVMGYOKRLNSMDKTRHLISVADLFLVLTLPKAVADVAVN
YFGNLFCKAVHIVITVNLVSSVLLAFISLDRLAIVAHATNSQRPRLAKVYVYG
WIRPALLTIRHFRITASVSEADDRICDRFYNDLVVYVFOHIMVGLIPGIVILSC
YCLITSLSHSGKQKRKALKTVILILAFACWLPYIGISIDFSIFLLEIKQCE
ENTVHKWISTTEALAFPHCCNPLIYLAFLGAKFKTSAOHALTSVRSGLSKILSKGR
GSHSSVSTSESSSFHSS"

CDS

BASE COUNT 241 a 300 c 237 g 290 t
ORIGIN

Query Match 7.5%; Score 131; DB 9; Length 1068;
Best Local Similarity 99.5%; Pred. No. 4.6e-60;
Matches 181: Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 631 GATGACAGTATATGCTGACCGCTTCTACCCCAATGACTTGGTGTGTGTTCCAG 690
DB 541 GATGACAGTATATGCTGACCGCTTCTACCCCAATGACTTGGTGTGTGTTCCAA 600
OY 691 TTTCAGCATATGATGGTGGCTTATCCCTGCTGATGTATGCTGCTGATATTC 750
DB 601 TTTCAGCATATGATGGTGGCTTATCCCTGCTGATGTATGCTGCTGATATTC 660
OY 751 ATTATCATCTCCAAAGCTGTACACTTCGAAGGCCACACGAGCGAAGGCCCTCAAGACC 810
DB 651 ATTATCATCTCCAAAGCTGTACACTTCGAAGGCCACACGAGCGAAGGCCCTCAAGACC 720
OY 811 AC 812
DB 721 AC 722

RESULT 24

LOCUS FCA9816 1689 bp mRNA MAM 03-MAY-1999
DEFINITION Felis catus mRNA for CXCR4 chemokine receptor.
ACCESSION FCA9816
NID 94200300
VERSION FCA9816.1 GI:4200300
KEYWORDS cell surface receptor; chemokine receptor; CXCR4 gene.
SOURCE cat.
ORGANISM Felis catus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Carnivora; Fissipedia; Felidae; Felis.

REFERENCE 1 (bases 1 to 1689)
AUTHORS Kovacs, E.M., Baxter, G.D. and Robinson, W.F.
TITLE Feline peripheral blood mononuclear cells express message for both CXC and CC type chemokines
JOURNAL Arch. Virol. 144: 273-285 (1999)
REFERENCE 2 (bases 1 to 1689)
AUTHORS Kovacs, E.M.
TITLE Direct Submission
JOURNAL Submitted (27-JUL-1998) Kovacs E.M., Department of Medicine, The Princess Alexandra Hospital, First Floor, Lions Research Unit, Ipswich Road, Woolloongabba, 4102, AUSTRALIA

FEATURES
SOURCE Location/Qualifiers
1..1689
/organism="Felis catus"
/db_xref="taxon:9685"
/sex="female"
/cell_line="primary lymphocyte"
/cell_type="peripheral blood mononuclear cells"
/dev_stage="adult"
70..1131
/gene="CXCR4"
70..1131
/gene="CXCR4"
/function="cell surface receptor"
/codon_start=1
/product="CXCR4 chemokine receptor"
/protein_id="CAA08839.1"
/db_xref="PID:el373135"
/db_xref="PID:94200301"
/db_xref="GI:4200301"
/translation="MDGFRITPSNYTDEDLGSGDYDSMKKEPCFRENAHNRIFLPTI
VSIITFLTGIVNGLVILVMGYOKRLNSMDKTRHLISVADLFLVLTLPKAVADVAVN
WYEGKFLCKAVHIVITVNLVSSVLLAFISLDRLAIVAHATNSQRPRLAKVYVYG
WIRPALLTIRHFRITASVSEADDRICDRFYNDLVVYVFOHIMVGLIPGIVILSC
YCLITSLSHSGKQKRKALKTVILILAFACWLPYIGISIDFSIFLLEIKQCE
FESTVHKWISTTEALAFPHCCNPLIYLAFLGAKFKTSAOHALTSVRSGLSKILSKGR
RGHSSVSTSESSSFHSS"

CDS

BASE COUNT 427 a 382 c 357 g 523 t
ORIGIN

Query Match 4.4%; Score 77; DB 3; Length 1689;
Best Local Similarity 100.0%; Pred. No. 7.6e-31;

Matches 77: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 415 TCGAAGCAGTCATGTCATCTACACAGTCACAGCTGACAGCTGACCTGCTGCC 474
|||||
Db 397 TCGAAGCAGTCATGTCATCTACACAGTCACAGCTGACAGCTGCTGCTGCC 456
|||||

QY 475 TTCATCAGCTGACCG 491
|||||

Db 457 TTCATCAGCTGACCG 473
|||||

RESULT 25
FCU63558 1170 bp mRNA MAM 06-APR-1998
LOCUS Felis catus CXCR-4 homolog mRNA, complete cds.
DEFINITION
ACCESSION U63558
NID 93025443
VERSION 063558.1 GI:3025443
KEYWORDS
SOURCE .
ORGANISM cat.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Carnivora; Fissipedia; Felidae; Felis.
REFERENCE 1 (bases 1 to 1170)
AUTHORS Willett,B.J., Picard,L., Hosie,M.J., Turner,J.D., Adema,K. and Clapham,P.R.
TITLE Shared usage of the chemokine receptor CXCR4 by the feline and human immunodeficiency viruses
J. Virol. 71 (9), 6407-6415 (1997)

JOURNAL MEDLINE 97404646
REFERENCE 2 (bases 1 to 1170)
AUTHORS Willett,B.J.
TITLE Direct Submission
JOURNAL Submitted (11-JUL-1996) Veterinary Pathology, University of Glasgow, Bearsden Road, Glasgow G61 1QH, UK
REFERENCE 3 (bases 1 to 1170)
AUTHORS Willett,B.J.
TITLE Direct Submission
JOURNAL Submitted (06-APR-1998) Veterinary Pathology, University of Glasgow, Bearsden Road, Glasgow G61 1QH, UK
REMARK Sequence update by submitter
COMMENT On Apr 6, 1998 this sequence version replaced gi:1731648.
FEATURES
source
Location/Qualifiers
1..1170
/organism="Felis catus"
/strain="domesticus"
/db_xref="taxon:9685"
/cell_type="T lymphoblasts"
84..1145
/function="receptor for chemokine SDF-1"
/note="LESTR"
/codon_start=1
/product="CXCR-4 homolog"
/protein_id="AAC48852.1"
/db_xref="PID:g1842247"
/db_xref="GI:1842247"
/translation="MDGFRIPSDNYTEDDLGSGDYDSKMEPCFRENAHFNRIPLPT
VYSIIFLTGIVNGVLIVMGYKLRSMIDKRLHSVDLFLVTLFPWADVAVAN
WYRKFELCKAVHYITVNISSVLIASFISDRYLAIIVAHATNSQRPRLIAEKVYVG
VWIPALLITIPDFIFANVREADGRITCDRFPSSSWLVVFOFQIMGLIPGVILS
CYCIIISKLSHSGYOKRKALKTVIILAFACWLPYIGISIDSLILEITKOGCE
FESTVHKWISITLALAFPHCCINPILVAFLGAKFKTSAOHALISVSRGSSLSKILSKG
RGHSSVSTESSESSFRHS"

BASE COUNT 268 a 320 c 269 g 313 t

ORIGIN

Query Match 4.4%; Score 77; DB 3; Length 1170;
Best Local Similarity 100.0%; Pred. No. 7.7e-31;
Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 415 TCGAAGCAGTCATGTCATCTACACAGTCACAGCTGACAGCTGCTGCC 474
|||||

Db 411 TCGAAGCAGTCATGTCATCTACACAGTCACAGCTGACAGCTGCTGCC 470

QY 475 TTCATCAGCTGACCG 491
|||||

Db 471 TTCATCAGCTGACCG 487
|||||

RESULT 26
FCU92795 1140 bp mRNA MAM 14-APR-1997
LOCUS Felis catus fusin (CXCR4) mRNA, complete cds.
DEFINITION
ACCESSION U92795
NID 91935044
VERSION 092795.1 GI:1935044
KEYWORDS
SOURCE .
ORGANISM cat.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Carnivora; Fissipedia; Felidae; Felis.
REFERENCE 1 (bases 1 to 1140)
AUTHORS Lerner,D.L. and Elder,J.H.
TITLE Feline CXCR4
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1140)
AUTHORS Lerner,D.L. and Elder,J.H.
TITLE Direct Submission
JOURNAL Submitted (10-MAR-1997) Molecular Biology, TSRI, 10666 North Torrey
Pines Road, La Jolla, CA 92037, USA
FEATURES
source
Location/Qualifiers
1..1140
/organism="Felis catus"
/db_xref="taxon:9685"
/cell_line="MCH5-4"
/clone="15"
/note="lymphocytic cell line"
1..1062
/gene="CXCR4"
1..1062
/gene="CXCR4"
/note="CXCR4"
/note="LESTR"
/codon_start=1
/product="fusin"
/protein_id="AAB51765.1"
/db_xref="PID:g1935045"
/db_xref="GI:1935045"
/translation="MDGFRIPSDNYTEDDLGSGDYDSKMEPCFRENAHFNRIPLPT
VYSIIFLTGIVNGVLIVMGYKLRSMIDKRLHSVDLFLVTLFPWADVAVAN
WYRKFELCKAVHYITVNISSVLIASFISDRYLAIIVAHATNSQRPRLIAEKVYVG
VWIPALLITIPDFIFANVREADGRITCDRFPSSSWLVVFOFQIMGLIPGVILS
CYCIIISKLSHSGYOKRKALKTVIILAFACWLPYIGISIDSLILEITKOGCE
FESTVHKWISITLALAFPHCCINPILVAFLGAKFKTSAOHALISVSRGSSLSKILSKG
RGHSSVSTESSESSFRHS"

BASE COUNT 257 a 312 c 255 g 314 t 2 others

ORIGIN

Query Match 4.4%; Score 77; DB 3; Length 1140;
Best Local Similarity 100.0%; Pred. No. 7.7e-31;
Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 415 TCGAAGCAGTCATGTCATCTACACAGTCACAGCTGACAGCTGCTGCC 474
|||||

Db 328 TCGAAGCAGTCATGTCATCTACACAGTCACAGCTGACAGCTGCTGCC 387
|||||

QY 475 TTCATCAGCTGACCG 491
|||||

Db 388 TTCATCAGCTGACCG 404
|||||

RESULT 27
HUMNRY92 95 bp mRNA PRI 07-FEB-1999
LOCUS Human mRNA for neuropeptide Y3 receptor, 5'UTR (sequence from the
DEFINITION

Accession	5' cap to the start codon.
NID D28433	
VERSION 9461211	
KEYWORDS D28433.1 GI:461211	
SOURCE neuropeptide Y3 receptor.	
ORGANISM Homo sapiens (library: U937/pKAL) lymphoma cell-line U937 CDNA to mRNA, clone HP00400.	
REFERENCE Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.	
AUTHORS Kato, S.	
TITLE 1 (bases 1 to 95)	
JOURNAL Direct Submission	
REFERENCE Submitted (03-FEB-1994) to the DDBJ/EMBL/GenBank databases. Seishi Kato, Sagami Chemical Research Center, Genetic Engineering Section 4-4-1, Nishi-Ohnuma, Sagamihara, Kanagawa 229, Japan (E-mail: bdn00121@biotech.com, Tel:0427-42-4791(ex.415), Fax:0427-49-7631)	
AUTHORS 2 (bases 1 to 95)	
JOURNAL Unpublished (1994)	
REFERENCE 3 (sites)	
AUTHORS Kato, S., Sekine, S., Oh, S.W., Kim, N.S., Umezawa, Y., Abe, N., Yokoyama-Kobayashi, M. and Aoki, T.	
TITLE Construction of a human full-length CDNA bank	
JOURNAL Gene 150 (2), 243-250 (1994)	
MEDLINE 95121910	
COMMENT Submitted (03-Feb-1994) to DDBJ by: Seishi Kato	
Genetic Engineering Section	
Sagami Chemical Research Center	
4-4-1 Nishi-Ohnuma	
Sagamihara, Kanagawa 229	
Japan	
Phone: 0427-42-4791	
Fax: 0427-49-7631	
FEATURES	
source	
1..95	
Location/Qualifiers	
/organism="Homo sapiens"	
/db_xref="taxon:9606"	
/cell_line="U937"	
/clone_lib="U937/pKAL"	
/tissue_type="lymphoma"	
10..89	
90..>95	
/codon_start=1	
/product="neuropeptide Y3 receptor"	
/protein_id="AA05798.1"	
/db_xref="PID:d1006346"	
/db_xref="PID:g4433177"	
/db_xref="GI:4433177"	
/translation="ME"	
BASE COUNT 21 a 24 c 33 g 17 t	
ORIGIN	
Query Match 4.0%; Score 70; DB 9; Length 95;	
Best Local Similarity 100.0%; Pident. No. 5.2e-27;	
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY 27 GCAAGTAGCCCGGAGGCGCTGAGTGCACGATAGCCACCGCATCTGGAGACCAAGCGT 86	
DB 26 GCAAGTAGCCCGGAGGCGCTGAGTGCACGATAGCCACCGCATCTGGAGACCAAGCGT 85	
QY 87 TACCATGGAG 96	
DB 86 TACCATGGAG 95	
RESULT 28	
LOCUS BOVNYR 1280 bp mRNA	
DEFINITION Cow neuropeptide Y receptor mRNA sequence.	
12-FEB-1992	

[illegible]

BASE COUNT 285 a 342 c 268 g 328 t
ORIGIN
TTFIFLITGIVNGVLIVLMGYOKLRKSTDKYRLHLVSADLLFYITLDPFNAVADMA
DMYFGKFLCKAVHIIYTVNLVSSVLIAFISLDRYLAIVHATNSQRPRLAEKAVY
GWMIPALLITIPDIFADVSQDISOGRDIDRLAYPSLWVVOFOHIMGLLP
GIVILSCYCIILIKLSHSGHOKRKALKTIVILIAFFACMLPYVVGISIDSPILIGV
IKOGCPESIVHKWISTITELAFHCCLNPIIAFLGAFAKSSAQAHLNMSRSGSLK
ILSKRGHSGSVSTSESSSFHSS"

Query Match 3.2%: Score 56; DB 12; Length 1223;
Best Local Similarity 100.0%; Pred. No. 1.8e-19;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 754 ATCATCTCCAGCTGTCCACACTCCAGGCCACAGAGGCCCTCAAGAC 809
DB 714 ATCATCTCCAGCTGTCCACACTCCAGGCCACAGAGGCCCTCAAGAC 769

RESULT 30
LOCUS MMU65580 3366 bp DNA ROD 14-DEC-1996
DEFINITION Mus musculus fusin (CXCR-4) gene, complete cds.
ACCESSION U65580
NID 91731650
VERSION U65580.1 GI:1731650
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS Heesen, M., Berman, M.A., Benson, J.D., Gerard, C. and Dorf, M.E.
TITLE Cloning of the mouse fusin gene, homologue to a human HIV-1
co-factor.

JOURNAL J. Immunol. 157 (12), 5455-5460 (1996)
MEDLINE 9711334
REFERENCE 2 (bases 1 to 3366)
AUTHORS Heesen, M., Berman, M.A., Gerard, C. and Dorf, M.E.
TITLE Direct Submission
JOURNAL Submitted (30-JUL-1996) Pathology, Harvard Medical School, 200
Longwood Avenue, Boston, MA 02115, USA

FEATURES
source
1. .3366
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
1. .21
/gene="CXCR-4"
/number=1
1. .15
/gene="CXCR-4"
/note="alternate exon 1"
1. .3366
/gene="CXCR-4"
join(1..21,2308..3366)
/gene="CXCR-4"
/note="SDF-1 chemokine receptor"
/codon_start=1
/product="fusin"
/protein_id="AAC52953.1"
/db_xref="PID:g3645915"
/translation="MEPIVSITSDNYSEVSGSDYDSNKEPCFRDENENRIFLPL
TIFYIFLITGIVNGVLIVLMGYOKLRKSTDKYRLHLVSADLLFYITLDPFNAVADMA
DMYFGKFLCKAVHIIYTVNLVSSVLIAFISLDRYLAIVHATNSQRPRLAEKAVY
GWMIPALLITIPDIFADVSQDISOGRDIDRLAYPSLWVVOFOHIMGLLP
GIVILSCYCIILIKLSHSGHOKRKALKTIVILIAFFACMLPYVVGISIDSPILIGV
IKOGCPESIVHKWISTITELAFHCCLNPIIAFLGAFAKSSAQAHLNMSRSGSLK
ILSKRGHSGSVSTSESSSFHSS"
16. >2307
/gene="CXCR-4"
/note="alternate intron 1"

Intron

intron 22. .2307
/gene="CXCR-4"
/number=1
exon 2308. .3366
/gene="CXCR-4"
/number=2
BASE COUNT 700 a 874 c 849 g 943 t
ORIGIN

Query Match 3.2%: Score 56; DB 12; Length 3366;
Best Local Similarity 100.0%; Pred. No. 1.7e-19;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 754 ATCATCTCCAGCTGTCCACACTCCAGGCCACAGAGGCCCTCAAGAC 809
DB 2971 ATCATCTCCAGCTGTCCACACTCCAGGCCACAGAGGCCCTCAAGAC 3026

RESULT 31
LOCUS RNU54791 1362 bp mRNA ROD 05-JUN-1996
DEFINITION Rattus norvegicus chemokine receptor LCR1 mRNA, complete cds.
ACCESSION U54791
NID 91354504
VERSION U54791.1 GI:1354504
KEYWORDS
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE
AUTHORS Wong, M., Xin, W.W. and Duman, R.S.
TITLE Rat LCR1: Cloning and cellular distribution of a putative chemokine
receptor in brain
JOURNAL Molecular Psychiatry (1996) In press
REFERENCE 2 (bases 1 to 1362)
AUTHORS Wong, M.

JOURNAL Direct Submission
TITLE Submitted (11-APR-1996) Ma-Li Wong, Clinical Neuroendocrinology
Branch, National Institutes of Health, NIMH, Building 10 Room
3S231, 10 Center DR MSC 1284, Bethesda, MD 20892-1284, USA

FEATURES
source
1. .1362
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="RLCR1"
44. .1042
/codon_start=1
/product="chemokine receptor LCR1"
/protein_id="AAB01981.1"
/db_xref="PID:g1354505"
/db_xref="GI:1354505"
/translation="WEITSDNYSEVSGSDYDSNKEPCFRDENENRIFLPTIFYI
IFLTGIVNGVLIVLMGYOKLRKSTDKYRLHLVSADLLFYITLDPFNAVADMA
DMYFGKFLCKAVHIIYTVNLVSSVLIAFISLDRYLAIVHATNSQRPRLAEKAVY
GWMIPALLITIPDIFADVSQDISOGRDIDRLAYPSLWVVOFOHIMGLLP
GIVILSCYCIILIKLSHSGHOKRKALKTIVILIAFFACMLPYVVGISIDSPILIGV
IKOGCPESIVHKWISTITELAFHCCLNPIIAFLGAFAKSSAQAHLNMSRSGSLK
ILSKRGHSGSVSTSESSSFHSS"

CDS

BASE COUNT 316 a 346 c 320 g 380 t
ORIGIN

Query Match 3.2%: Score 56; DB 12; Length 1362;
Best Local Similarity 100.0%; Pred. No. 1.8e-19;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 754 ATCATCTCCAGCTGTCCACACTCCAGGCCACAGAGGCCCTCAAGAC 809
DB 647 ATCATCTCCAGCTGTCCACACTCCAGGCCACAGAGGCCCTCAAGAC 702

```

RESULT 32
RN090610 1050 bp mRNA ROD 26-MAR-1997
LOCUS Rattus norvegicus CXc chemokine receptor (CXCR4) mRNA, complete
DEFINITION
cds.
ACCESSION U90610
NID 91906612
VERSION U90610.1 GI:1906612
KEYWORDS
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
1 (bases 1 to 1050)
REFERENCE 1
AUTHORS Harrison, J.K. and Salafianca, M.N.
TITLE Molecular cloning of rat CXCR4
JOURNAL Unpublished
2 (bases 1 to 1050)
REFERENCE 2
AUTHORS Harrison, J.K. and Salafianca, M.N.
TITLE Direct Submission
JOURNAL Submitted (25-FEB-1997) Pharmacology and Therapeutics, University
of Florida, 1600 SW Archer Rd., Gainesville, FL 32610-0267, USA
FEATURES
Source
Location/Qualifiers
1..1050
/organism="Rattus norvegicus"
/strain="Wistar"
/db_xref="taxon:10116"
/tissue_type="spleen"
1..1050
/gene="CXCR4"
1..1050
/gene="CXCR4"
/feature="G-protein coupled receptor"
/codon_start=1
/product="CXC chemokine receptor"
/protein_id="AAB50408.1"
/db_xref="PID:91906613"
/db_xref="GI:1906613"
/translacion="MEITYSDNYSSEVSGDYDSNKEPCFDENMENRIFLPTIFY
ILFLIGVNGCIVLVMGYOKRLRSMTOGRLSLVADLIPFITTLPFAVDMAAMWYFS
KFLCKAVHIITVNLVSSVLLAFISDRITATVHATNSOSARLLEKAVYGVWIP
ALLLTIPDIIPADVSQDGRICDRLVPSLMMVVFQHLVGLILPGIVILCYCI
IISKLSHKQKRRKALKTIVILIAFFACWLPYVGISIDSFILLEVIKGCFFESY
VHKWISITELAFHCCINPLIYAFILGAKFKSSAQAHLNMSRSGSLKILSKRGKH
SSVSTESSESSFHSS"
BASE COUNT 231 a 313 c 247 g 259 t
ORIGIN
Query Match 3.2%; Score 56; DB 12; Length 1050;
Best Local Similarity 100.0%; Pred. No. 1.8e-19;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 754 ATCATCTCAAGCTGTCTCACTCCAGAGCCACGAAAGCGCAGCCCTCAAGAC 809
Db 655 ATCATCTCAAGCTGTCTCACTCCAGAGCCACGAAAGCGCAGCCCTCAAGAC 710
RESULT 33
LOCUS E07381 3581 bp RNA PAT 29-SEP-1997
DEFINITION cDNA encoding mouse OSF-4.
ACCESSION E07381
NID 92175520
VERSION E07381.1 GI:2175520
KEYWORDS JP 1994122700-A/1.
SOURCE Mus musculus.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 3581)
REFERENCE 1
AUTHORS Takeshita, A., Okazaki, M., Kawai, S., Tsujimura, A. and Aman, E.
TITLE BONE-RELATED CADHERIN-LIKE PROTEIN AND ITS PRODUCTION

```

```

JOURNAL Patient: JP 1994122700-A 1 06-MAY-1994;
COMMENT HOECHST JAPAN LTD
OS Mus musculus (mouse)
PN JP 1994122700-A/1
PD 06-MAY-1994
PE 13-JUL-1993 JP 1993172883
PR 28-AUG-1992 JP 92P 230028
PI TAKESHITA ATSUSHI, OKAZAKI MAKOTO, KAWAI SHINJI, PI
TSUTSUMI ATSUSHI,
PI AMAN EGON
PC C07K13/00,A61K37/02,A61K39/395,A61K49/00,
PC C07K15/14,
PC C12N15/12,C12P21/02,C12P21/08;
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FH source 1..3581
FT /organism="Mus musculus"
FT /strain="MC3T3-E1"
FT /clone="pKOT164"
FT 5'UTR 1..283
FT CDS 284..2674
FT /product="Mouse OSF-4"
FT 2675..3581.
FT 3'UTR
Location/Qualifiers
1..3581
/organism="Mus musculus"
/db_xref="taxon:10090"
BASE COUNT 1121 a 773 c 801 g 886 t
ORIGIN
Query Match 3.1%; Score 53; DB 5; Length 3581;
Best Local Similarity 100.0%; Pred. No. 7.2e-18;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1669 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCC 1721
Db 3523 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCC 3575
RESULT 34
LOCUS AF002985 995 bp mRNA PRI 04-MAR-1999
DEFINITION Homo sapiens putative alpha chemokine (H174) mRNA, complete cds.
ACCESSION AF002985
NID 92580585
VERSION AF002985.1 GI:2580585
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 995)
REFERENCE 1
AUTHORS Jacobs, K.A., Collins-Racie, L.A., Colbert, M., Duckett, M.,
Golden-Fleet, M., Kelleher, K., Kriz, R., Lavallee, E.R., Merberg, D.,
Spaulding, V., Stover, D., Williamson, M.J. and McCoy, J.M.
A genetic selection for isolating cDNAs encoding secreted proteins
Gene 198 (1-2), 289-296 (1997)
MEDLINE 98036061
JOURNAL 2 (bases 1 to 995)
REFERENCE 2
AUTHORS Luo, Y., Kim, R., Gabuzda, D., Mi, S., Collins-Racie, L.A., Lu, Z.,
Jacobs, K.A. and Dorf, M.E.
The CXc-chemokine, H174: expression in the central nervous system
J. Neurovirol. 4 (6), 575-585 (1998)
TITLE 99163782
JOURNAL 3 (bases 1 to 995)
REFERENCE 3
AUTHORS Jacobs, K.A., Collins-Racie, L.A., Colbert, M., Duckett, M.,
Golden-Fleet, M., Kelleher, K., Kriz, R., Lavallee, E.R., Merberg, D.,
Spaulding, V., Stover, D., Williamson, M.J. and McCoy, J.M.

```

TITLE Direct Submission
JOURNAL Submitted (07-MAY-1997) Genetics Institute, 87 Cambridge Park
Drive, Cambridge, MA 02140, USA

FEATURES
Source location/Qualifiers
1..995

gene
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="PHA and PMA activated human peripheral blood
mononuclear cells"
1..995
/gene="H174"
88..372
/gene="H174"
/codon_start=1
/product="putative alpha chemokine"
/protein_id="AAC51845.1"
/db_xref="PID:92580586"
/db_xref="GI:2580586"
/translation="MSVKGMAIALAVILCATVVGFPMEKRGRCICIGPVKAVVAD
IEKASIMPSNNCDIEVITLTKENKGGRCINPKSKQARLIIKKVERKNF"

CDS

BASE COUNT 382 a 170 c 194 g 249 t
ORIGIN

Query Match 3.1%; Score 53; DB 42; Length 995;
Best Local Similarity 100.0%; Pred. No. 7.6e-18;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1659 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCCG 1721
Db 943 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCCG 995

RESULT 35

LOCUS OAU38942 578 bp mRNA MAM 14-NOV-1995
DEFINITION Ovis aries neuropeptide Y receptor type 3 mRNA, partial cds.
ACCESSION U38942
NID g1061411
VERSION U38942.1 GI:1061411
KEYWORDS
SOURCE Sheep.
ORGANISM Ovis aries
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae;
Caprinae; Ovis.
1 (bases 1 to 578)
Dyer,C.J., Matteri,R.L. and Keisler,D.H.
Development of an ovine Y3 cDNA and expression of the Y3 receptor
mRNA in the ovine hypothalamus and pituitary
Abstr. Soc. Neurosci. 21, 1890-1890 (1995)
2 (bases 1 to 578)
Dyer,C.J., Matteri,R.L. and Keisler,D.H.
Direct Submission
Submitted (19-OCT-1995) Cheryl J. Dyer, Animal Science, University
of Missouri, 160 ASCC, Columbia, MO 65211, USA
Location/Qualifiers
1..578
/organism="Ovis aries"
/db_xref="taxon:9940"
/tissue_type="hypothalamus"
/note="synonyms: ICRL, IESTR, Y3 receptor; putative
neuropeptide Y receptor"
/codon_start=3
/product="neuropeptide Y receptor type 3"
/protein_id="AA81347.1"
/db_xref="PID:g1061412"
/db_xref="GI:1061412"
/translation="YTDDLSGGDYDSKKEPCRPENAHENRIFLPTVYSIIPLGIV
GNGILVIMGOKLRSMSTDKYRLHLSTADLLFLILPFAVADAVANWYRPGFLCKAV
HVIITVLISSVLLAFISDRLTAIYHAITSQPRKLLAEKVVYGVWLPVALLTLP
DLIADIKAEADERICDFIPSDLWLVFQFQ"

REFERENCE
AUTHORS
TITLE
JOURNAL
AUTHORS
REFERENCE
TITLE
JOURNAL
FEATURES
Source

CDS

BASE COUNT 121 a 157 c 146 g 154 t
ORIGIN

Query Match 2.9%; Score 50; DB 3; Length 578;
Best Local Similarity 100.0%; Pred. No. 3.7e-16;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 442 GTCAACCTCTACAGCAGTGTCTCATCTGGCTTCATCAGTGTGACCG 491
Db 324 GTCAACCTCTACAGCAGTGTCTCATCTGGCTTCATCAGTGTGACCG 373

RESULT 36

LOCUS I88893 60 bp DNA PAT 17-JUL-1998
DEFINITION Sequence 11 from patent US 5719125.
ACCESSION I88893
NID g3408833
VERSION I88893.1 GI:3408833
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
Source

Query Match 2.9%; Score 50; DB 5; Length 60;
Best Local Similarity 100.0%; Pred. No. 3.5e-16;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1674 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCCA 1723
Db 60 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCCA 11

RESULT 37

LOCUS D87747 1877 bp mRNA ROD 07-FEB-1999
DEFINITION Mus musculus mRNA for murine CXCR-4, complete cds.
ACCESSION D87747
NID g1772445
VERSION D87747.1 GI:1772445
KEYWORDS
SOURCE murine CXCR-4.
ORGANISM Mus musculus
Mus musculus bone marrow pre-B cell cell_line:DW34 cDNA to mRNA.
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1877)
Nagasawa,T.
Direct Submission
Submitted (05-SEP-1996) to the DDBJ/EMBL/Genbank databases. Takashi
Nagasawa, Research Institute, Osaka Medical Center, Department of
Immunology; Murdockho 840, Izumi, Osaka 590-02, Japan
(E-mail:immunol@osk.threewebnet.or.jp, Tel:0725-561220,
Fax:0725-57-3021)
2 (bases 1 to 1877)
Nagasawa,T.
Molecular cloning of murine PBSF/SDF-1 receptor
Unpublished (1996)
3 (sites)
Nagasawa,T., Nakajima,T., Tachibana,K., Iizasa,H., Bleul,C.C.,
Yoshie,O., Matsushima,K., Yoshida,N., Springer,T.A. and
Kishimoto,T.

REFERENCE
AUTHORS
TITLE
JOURNAL
AUTHORS
REFERENCE
TITLE
JOURNAL
FEATURES
Source

TITLE Molecular cloning and characterization of a murine pre-B-cell growth-stimulating factor/stromal cell-derived factor-1 receptor, a murine homolog of the human immunodeficiency virus 1 entry coreceptor fusin

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 93 (25), 14726-14729 (1996)

MEDLINE 97121456

FEATURES

source location/Qualifiers

1..1877

/organism="Mus musculus"

/db_xref="taxon:10090"

/cell_line="DM34"

/cell_type="pre-B cell"

/tissue_type="bone marrow"

120..1199

/gene="murine CXCR-4"

120..1199

/gene="murine CXCR-4"

/function="PBSF/SDF-1 receptor"

/codon_start=1

/evidence="experimental"

/product="murine CXCR-4"

/protein_id="BA013451.1"

/db_xref="PID:d1014141"

/db_xref="PID:g1772446"

/db_xref="GI:1772446"

/translation="MEPISVSYTSDNYSFEGSGDYDSNKEPCFREDENYHFRILFPTIFETITGVGNGVILVIMGYOKRKRTMTDRLHLSTADLFTVTEPFAVDADA DMTEFGFLCAVHITVTNLYSSVLIATFSLDRYLAIVATNSORRKLAEAVIV GWIPALFLTPDFIFADVSODISODDDYCDRLYPSLMMVVFQFQIMVGLIIP GIVILSCYCIHISKLHSGHOKRKALITVILIAFFACMLPYVGVISIDSFILGV IKGCGDFESIVHKWISITELAFHCCINPILYAFILGAKFKSSAOHLNMSGSSLIK ILSKGRGHSSTSESSSFSS"

BASE COUNT 469 a 450 c 407 g 551 t

ORIGIN

Query Match 2.9%; Score 50; DB 12; Length 1877; Best Local Similarity 100.0%; Pred. No. 3.1e-16; Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1662 TAAACTTAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1711

DB 1828 TAAACTTAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1877

RESULT 38

LOCUS S81785 495 bp mRNA INV 21-JUL-1996

DEFINITION Myr p II=major allergen [Myrmecia pilosula=Australian jumper ants, abdomens, mRNA Partial, 495 nt].

ACCESSION S81785

NID g1438760

VERSION S81785.1 GI:1438760

KEYWORDS

SOURCE Australian jumper ant abdomens.

ORGANISM Myrmecia pilosula

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Formicoidea; Formicidae; Myrmecia.

1 (bases 1 to 495)

REFERENCE

AUTHORS Street, M.D., Donovan, G.R. and Baldo, B.A.

TITLE Molecular cloning and characterization of the major allergen Myr p II from the venom of the jumper ant Myrmecia pilosula: Myr p I and Myr p II share a common protein leader sequence

JOURNAL Biochim. Biophys. Acta 1305 (1-2), 87-97 (1996)

MEDLINE 96180991

REMARK GenBank staff at the National Library of Medicine created this entry [NCBI g143876] from the original journal article. This sequence comes from Fig. 5.

FEATURES

1..495 location/Qualifiers

/organism="Myrmecia pilosula"

/db_xref="taxon:13618"

gene 44..271

/gene="Myr p II"

CDS 44..271

/gene="Myr p II"

/note="major allergen; Method: conceptual translation with partial peptide sequencing. This sequence comes from Fig. 5"

/codon_start=1

/product="Myr p II"

/protein_id="AAB36316.1"

/db_xref="PID:g1438761"

/db_xref="GI:1438761"

/translation="MKISCLLLALIFVITIVHAPVVEAKALADPESDAVGFADAVG EADPIDMKVDMKVSKRKTCKVLMKAKKFLG"

BASE COUNT 201 a 68 c 99 g 127 t

ORIGIN

Query Match 2.9%; Score 50; DB 35; Length 495; Best Local Similarity 100.0%; Pred. No. 3.3e-16; Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1672 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCCGC 1721

DB 446 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCCGC 495

RESULT 39

LOCUS AF139986 3362 bp mRNA VRT 30-MAY-1999

DEFINITION xenopus laevis cleavage and polyadenylation specificity factor 100 kDa subunit mRNA, complete cds.

ACCESSION AF139986

NID g4927239

VERSION AF139986.1 GI:4927239

KEYWORDS

SOURCE African clawed frog.

ORGANISM Xenopus laevis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae; Xenopodinae; Xenopus.

1 (bases 1 to 3362)

REFERENCE

AUTHORS Dickson, K.S., Bilger, A., Ballantyne, S. and Wickens, M.P.

TITLE CPSP in Xenopus laevis oocytes: a cytoplasmic factor involved in regulated polyadenylation

JOURNAL Unpublished

REFERENCE

AUTHORS Dickson, K.S. and Wickens, M.P.

TITLE Direct Submission

JOURNAL Submitted (01-APR-1999) Biochemistry, UW-Madison, 433 Babcock Dr., Madison, WI 53706, USA

FEATURES

1..3362 location/Qualifiers

/organism="Xenopus laevis"

/db_xref="taxon:8355"

143..2494

/note="CPSP"

/codon_start=1

/product="cleavage and polyadenylation specificity factor 100 kDa subunit"

/protein_id="AAB33061.1"

/db_xref="PID:g4927240"

/db_xref="GI:4927240"

/translation="MTSIKLTITVGAGEESAVCYLQYDEPRFLIDCGMDENFSMDI IDSVKYVQVAVLISHDPLHLGLPFAVNGLGNCAIYATIPYKKGOMPMYLY OSKRNTEDSLSDVDCAFDKIQDLKNOVHLKGGHGSITLPPLPGHMTGIIW KLYKDEEELIYAVDNRKREHLNCSLEMINRPLLTDSFNATYVOPRRKQROEQ LTNVLETLRGQNVLIADVIAGVLELQOLDQWRKTADAGVSYSLANVSYNV VESKSOVEMMSDKLMRCFEDKRNPFORHLITLCHGYSDLAVPSPKVLASOPLE CGFSRELFIQCODPKNSVILYRTPTGLARFLIDPESRIIDIELRRKRLKEGKL EYVEKERLKTFAKRLKOSKEADDSDDSVVEDIDITSKAKHDLMKRNESRK GSFFKQAKSYPPFPAPEDRIRKWDYGEIIRKPEDFLVPELOVTEDEKTLUESQLNGD EPMDDLSVPTKCVSTESMEIKARVITYIDEGRSDGSIKKIINOMKPROLIIVHG


```

Query Match      2.8%; Score 49; DB 8; Length 1207;
Best Local Similarity 100.0%; Pred. No. 1.1e-15;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1665 ACTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1713
|||||
Db 1148 ACTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1196

RESULT 44
HSM800467 HSM800467 1845 bp mRNA PRI 21-MAY-1999
LOCUS Homo sapiens mRNA; CDNA DKFZps586f1322 (from clone DKFZps586f1322).
ACCESSION AL050172
NID 94884384
VERSION AL050172.1 GI:4884384
KEYWORDS
SOURCE
ORGANISM
human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 1845)
Koehler,K., Beyer,A., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
Direct Submission
Submitted (21-MAY-1999) MIPS, Am Klopfer spitzz 18a D-82152
Martinsried, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); sequenced by BHFZ within the CDNA
sequencing consortium of the German Genome Project. This clone is
available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
Source
location/Qualifiers
1..1845
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZps586f1322"
/clone_id="586 (synonym: hutel). Vector pSport1; host
DH10B; sites NotI + SalI/MluI"
/dev_stage="adult"
/tissue_type="uterus"
1743..1748
polyA_site
BASE COUNT 521 a 365 c 356 g 603 t
ORIGIN

Query Match      2.8%; Score 49; DB 10; Length 1845;
Best Local Similarity 100.0%; Pred. No. 1.1e-15;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1665 ACTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1713
|||||
Db 1759 ACTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1807

RESULT 45
MMU132433 MMU132433 1403 bp mRNA ROD 27-MAY-1999
LOCUS Mus musculus untranslated RNA G90.
DEFINITION AJ132433
ACCESSION g4914645
NID
VERSION AJ132433.1 GI:4914645
```

KEYWORDS G90.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Krause, R., Hemberger, M., Himmelbauer, H., Kalschauer, V. and Fundele, R. H.

TITLE Identification and characterization of G90, a novel mouse RNA that lacks an extensive open reading frame

JOURNAL Gene 232 (1), 35-42 (1999)

MEDLINE 99267407
REFERENCE 2 (bases 1 to 1403)
AUTHORS Fundele, R. H.
TITLE Direct Submission
JOURNAL Submitted (20-JAN-1999) Fundele R. H., Rogers, Max-Planck-Institut fuer molekulare Genetik, Ihnestrasse 73, D-14195 Berlin, GERMANY

FEATURES
Source
Location/Qualifiers
1..1403
/organism="Mus musculus"
/strain="C57BL/6"
/db_xref="taxon:10090"
/chromosome="6"
/dev_stage="adult"
/cell_type="small intestine"
1..1357
/note="untranslated RNA G90"
1358

misc_RNA
polyA_site 392 a 336 c 367 g 306 t 2 others
BASE COUNT 392 a 336 c 367 g 306 t 2 others
ORIGIN

Query Match 2.8%; Score 49; DB 12; Length 1403;
Best Local Similarity 100.0%; Pred. No. 1.1e-15;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1665 ACTTAAA 1713
|||||
Db 1355 ACTTAAA 1403

RESULT 46
MUSCALTRA 754 bp mRNA ROD 04-FEB-1999
LOCUS Mouse mRNA for caltactin.
DEFINITION D16301
ACCESSION g467515
NID D16301.1 GI:467515
VERSION caltactin.
KEYWORDS Mus musculus (sub-species: domesticus, strain: BALB/c) testis cDNA to mRNA.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 754)
AUTHORS Ogawa, K.
TITLE Direct Submission
JOURNAL Submitted (20-MAY-1993) to the DBJ/EMBL/GenBank databases. Kazuo Ogawa, National Institute for Basic Biology, 38 Nishigonaka, Myodaiji-cho, Okazaki, Aichi 444, Japan (E-mail:ogawa@nibb.ac.jp, Tel:0564-53-7504, Fax:0564-53-7400)
2 (bases 1 to 754)
Ogawa, K. and Shimizu, T.
TITLE cDNA sequence for mouse caltactin
JOURNAL Biochim. Biophys. Acta 1216 (1), 126-128 (1993)
MEDLINE 94032472
COMMENT Submitted (20-MAY-1993) to DDBJ by: Kazuo Ogawa
Department of Cell Biology
National Institute for Basic Biology
38 Nishigonaka, Myodaiji-cho
Okazaki, Aichi 444
Japan

Phone: 0564-53-7504
Email: ogawa@nibb.ac.jp
Fax: 0564-53-7400
Location/Qualifiers
1..754
/organism="Mus musculus"
/strain="BALB/c"
/sub-species="domesticus"
/db_xref="taxon:10090"
/tissue_type="testis"
35..553
/codon_start=1
/product="caltactin"
/protein_id="BA03806.1"
/db_xref="PID:d1004322"
/db_xref="PID:g565281"
/db_xref="GI:565281"
/translation="MASTPRKSNVYASNYKRVGPPPELPEDEKQEVREAFDLPSDGSGTIDYKELKAMALGFEPKREMKKISVDGEATGKISFNFLVIMQKAKKPTKEETLKARLPDDETGKISFKNLKRVANLGEGLTDELDQEMIDEADRDGDGEVNEEFLKIMKKTMLY"

BASE COUNT 272 a 134 c 186 g 162 t
ORIGIN

Query Match 2.8%; Score 49; DB 12; Length 754;
Best Local Similarity 100.0%; Pred. No. 1.1e-15;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1665 ACTTAAA 1713
|||||
Db 684 ACTTAAA 732

RESULT 47
AFCATP 887 bp DNA INV 28-JUN-1994
LOCUS A. franciscana gene for sarco/endoplasmic reticulum Ca-ATPase.
DEFINITION X76991
ACCESSION g440382
NID X76991.1 GI:440382
VERSION Ca2+-ATPase.
KEYWORDS Artemia franciscana.
SOURCE Artemia franciscana.
ORGANISM Artemia franciscana.
Eukaryota; Metazoa; Arthropoda; Crustacea; Branchiopoda; Anostraca; Artemiidae; Artemia.
REFERENCE 1 (bases 1 to 887)
AUTHORS Escalante, R. and Sastre, L.
TITLE Structure of Artemia franciscana sarco/endoplasmic reticulum Ca-ATPase gene
JOURNAL J. Biol. Chem. 269 (17), 13005-13012 (1994)
MEDLINE 94230387
REFERENCE 2 (bases 1 to 887)
AUTHORS Sastre, L.
TITLE Direct Submission
JOURNAL Submitted (28-DEC-1993) L. Sastre, Inst. de Invest. Biomed. del CSIC, C/ Arturo Duperier, 4, 28029 Madrid, SPAIN

FEATURES
Source
Location/Qualifiers
1..887
/organism="Artemia franciscana"
/db_xref="taxon:6661"
/dev_stage="20 hours developed cysts"
/tissue_type="whole organism"
/clone_lib="EMBL3-genomic"
/clone="GAFATCa-1"
766..887
/product="Sarco/endoplasmic reticulum Ca-ATPase"
BASE COUNT 281 a 169 c 128 g 309 t
ORIGIN

Query Match 2.8%; Score 49; DB 36; Length 887;
Best Local Similarity 100.0%; Pred. No. 1.1e-15;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1666 CTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAG 1714
 Db 463 CTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAG 511

RESULT 48

LOCUS BMOTELRPT 562 bp DNA INV 26-FEB-1999
 DEFINITION Bombyx mori gene, telomeric repeat sequence.
 ACCESSION D13554
 NID 9217292
 VERSION D13554.1 GI:217292
 KEYWORDS telomeric repeat.
 SOURCE Bombyx mori (strain:Kinshu x Showa, haplotype:diploid) fifth larvae fat body DNA, clone_lib:lambd4 EMBL3 clone:pBluescript SK+.
 PB1-HK.

ORGANISM

Bombyx mori
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Lepidoptera; Bombycoidea; Bombycidae; Bombyx.

REFERENCE

1 (bases 1 to 562)
 Okazaki, S., Tsuchida, K., Maekawa, H., Ishikawa, H. and Fujiwara, H. Identification of a pentanucleotide telomeric sequence, (TTAGG)_n, in the silkworm Bombyx mori and in other insects

JOURNAL

Mol. Cell. Biol. 13 (3), 1424-1432 (1993)

MEDLINE

93180787

REFERENCE

2 (bases 1 to 562)

AUTHORS

Okazaki, S.

TITLE

Direct Submission

JOURNAL

Submitted (05-NOV-1992) to the DDBJ/EMBL/GenBank databases. Satoshi Okazaki, University of Tokyo, Faculty of Science, Zoological Institute, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113, Japan (Tel:03-2812-2111(ex.4449), Fax:03-3816-1965)

FEATURES

source

location/Qualifiers

1..562

/organism="Bombyx mori"

/strain="Kinshu x Showa"

/db_xref="taxon:7091"

/clone="pBluescript SK+", PB1-HK"

/clone_lib="lambd4 EMBL3"

/dev_stage="fifth larvae"

/haplotype="diploid"

/tissue_type="fat body"

/repeat_region

331..459

/rpt_family="telomeric repeat"

/rpt_type="direct"

BASE COUNT 122 a 83 c 134 g 223 t

ORIGIN

Query Match

Best Local Similarity 100.0%; Score 49; DB 36; Length 562;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1665 ACTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1713

Db 533 ACTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 485

RESULT 49

LOCUS AF013214 1032 bp mRNA MAM 04-AUG-1997
 DEFINITION Bos taurus acidic ribosomal phosphoprotein PO mRNA, partial cds.
 ACCESSION AF013214
 NID 92293576
 VERSION AF013214.1 GI:2293576
 KEYWORDS
 SOURCE Bos taurus.
 ORGANISM Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.

REFERENCE

1 (bases 1 to 1032)
 Lilienstiek, B., Rooha, M., Umansky, V., Benner, A., Lin, Y., Ziegler, R., Newroth, P.P. and Schirmacher, V.

TITLE

Direct Submission

Submitted (10-JUL-1997) Internal Medicine I, University Heidelberg, Im Neuenheimer Feld 324, 3. OG, R320, Heidelberg 69120, Germany

JOURNAL

location/Qualifiers

1..1032

/organism="Bos taurus"

/db_xref="taxon:9913"

/cell_type="endothelial"

/tissue_type="aorta"

/note="CDNA library CBAEC, enriched for TNF-alpha suppressed genes"

<1..910

/note="located within the 60S subunit of eukaryotic ribosomes; interacts with elongation factor 1 alpha and elongation factor 2 alpha; involved in protein synthesis"

/product="acidic ribosomal phosphoprotein PO"

/protein_id="AAB65436.1"

/db_xref="PIR:Q2293577"

/db_xref="GI:2293577"

/translation="IIQLDDYPCFIVGADNVGSKOMQOIRMSLGRKAVLVGNKNTM MKRAIRGHLENNPALKELPHIRGNVGFYTKEDLEIRDMLANVPAARAGALAP CEVTVPAAONTGLGPEKTSFEOALGITTKISRTGFEILSDPLKGTGDKGASATLLN MNISPSFGVIGVQVEVNGSTINPEVLTTEETJHSRFLGVRNVAACIOIGYPTV ASYPSHITNGKRYVLAISVETDVTFFPLAKRKAFLLDPSAFVAARVAAPAAAPAT TAAPAKVEAKKESESDMDGFELEPD"

BASE COUNT 283 a 271 c 264 g 214 t

ORIGIN

Query Match

Best Local Similarity 100.0%; Score 48; DB 3; Length 1032;

Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1666 CTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1713

Db 975 CTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1022

RESULT 50

LOCUS ECCD44 1338 bp mRNA MAM 07-SEP-1993
 DEFINITION E.caballus mRNA for lymphocyte surface antigen precursor CD44.
 ACCESSION X66862
 NID 91059
 VERSION X66862.1 GI:1059
 KEYWORDS lymphocyte surface antigen.
 SOURCE Equus caballus.
 ORGANISM Equus caballus

REFERENCE

1 (bases 1 to 1338)

Tavernor, A.A.

Direct Submission

Submitted (15-JUN-1992) A.S. Tavernor, Inst of Animal Physiology & Genetics, Research, Cambridge Station Babraham Hall, Cambridge CB2 4AT, UK

2 (bases 1 to 1338)

Tavernor, A.S., Deverson, E.V., Coadwell, W.J., Lunn, D.P., Zhang, C., Davis, W. and Butcher, G.W.

Molecular cloning of equine CD44 cDNA by a COS cell expression system

Immunogenetics 37 (6), 474-477 (1993)

JOURNAL 93170897

FEATURES

source

location/Qualifiers

1..1338

/organism="Equus caballus"

/isolate="MACD44"

/db_xref="taxon:9796"

/tissue_type="peripheral blood lymphocyte (PBL)"

sig_peptide
 CDS
 /cell_type="PBL 42 hour concanavalin A blasts"
 /clone_lib="CDNA"
 /clone="MACE044"
 55..114
 /product="lymphocyte surface antigen CD44"
 55..1134
 /codon_start=1
 /product="lymphocyte surface antigen precursor CD44"
 /protein_id="CAA47331.1"
 /db_xref="PID:g1060"
 /db_xref="GI:1060"
 /db_xref="SWISS-PROT:O05078"
 /translation="MDKFMWRAWGLCLVPLSLAQLDLNITCRAGVFHVEKNGRYSI
 SRTKADLCKARNSITPTMAQMKALNIGFETCRIGFISGHVYIPPIHNSICANN
 GYVILITSNISQDTYCFNASAPPEEDCTSYTDLPNAFEGPITITLYNRDGT
 YTKKE
 YRTNPEIDINPSTPADDDVSSSSSEKSTSGGYSIFTHLPITRPTODQSSPWYSDSPE
 KPTPTKDRASGGRAQTTHGSETSGHSTGSGGASTTSGPIRRPQIPFWLITLASILA
 LAILAVCIAVNSRRRCQKKLYINNGAVDRAKASGLNGEASRSQEWVHLVKNES
 SETQDOFMTADETRNLQNVDMKIGV"

mat_peptide 115..1131
 /product="lymphocyte surface antigen CD44"
 misc_feature 853..915
 /note="transmembrane region"
 polyA_signal 1249..1254
 polyA_site 1269..1329
 BASE COUNT 390 a 381 c 330 g 237 t
 ORIGIN

Query Match 2.8%; Score 48; DB 3; Length 1338;
 Best Local Similarity 100.0%; Pred. No. 3.8e-15;
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1666 CTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1713
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 1266 CTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1313

Search completed: August 7, 1999, 21:39:11
 Job time: 1141 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 8, 1999, 02:54:36 ; Search time 46.68 Seconds
(without alignments)
9309.847 Million cell updates/sec

Title: US-09-104-063-3

Perfect score: 1737

Sequence: 1 GAATTCAGTGTGCTGGCGG.....CCGCCAGCAGCTGGAATTC 1737

Scoring table: OLIGO_NUC

Searched: 311585 seqs, 125096042 residues

Database : N_Geneseq_36:*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1737	100.0	1737	1	080521 Human monocytic PF4
2	1737	100.0	1737	1	099007 Chemokine superfamily
3	1533	88.3	1737	1	029506 New platelet factor
4	1533	87.7	1944	1	X15882 CDNA encoding G-pr
5	1067	51.4	1317	1	V18357 Human PM3 seven tr
6	1016	58.5	1317	1	066179 Seven transmembran
7	249	14.3	611	1	X15883 CDNA encoding a pa
8	218	12.6	218	1	T20146 Human gene signal
9	53	3.1	3581	1	044391 Sequence of murine
10	53	3.1	1338	1	055607 Rabbit zona pelluc
11	53	3.1	87	1	V05740 Nucleotide sequenc
12	53	3.1	94	1	V05728 Nucleotide sequenc
13	53	3.1	4237	1	V61487 Human secreted pro
14	53	3.1	1338	1	V64789 Rabbit ZPC CDNA. I
15	53	3.1	2744	1	V84799 Nucleotide sequenc
16	51	2.9	61	1	078635 Human chondromodul
17	51	2.9	775	1	X26148 Human pan-s/tk-1A
18	51	2.9	1718	1	X26151 3' CDNA sequence o
19	50	2.9	495	1	T74484 Pilosulin 2 precu
20	50	2.9	201	1	V00418 3' fragment of clo
21	50	2.9	1877	1	V46370 Nucleic acid encod
22	49	2.8	56	1	052732 Sequence of oligo
23	49	2.8	55	1	052734 Rat glial cell der
24	49	2.8	2378	1	T84975 Secreted protein D
25	49	2.8	3443	1	V33197 Human membrane pro
26	49	2.8	2519	1	V54124 Protein PRO233 CN
27	49	2.8	1770	1	X52242 Rape acyl-ACP thio
28	48	2.8	1586	1	047967 Human superoxide d
29	48	2.8	699	1	058663 Gibberellin-20-ox
30	48	2.8	1490	1	075361 Thrombopoietin cod
31	48	2.8	861	1	098552 Oil seed rape cyst
32	48	2.8	1474	1	T90174 T. gondii chorisma
33	48	2.8	2312	1	V21646 Fanconi anaemia of
34	48	2.8	5503	1	V16187 Analytical solid p
35	48	2.8	90	1	V37789 Human secreted pro
36	48	2.8	1049	1	V43610 Coding sequence fo
37	48	2.8	772	1	V58363 Human secreted pro
38	48	2.8	1328	1	V58754 Human secreted pro
39	48	2.8	1373	1	V52969 Human G1 protein 1
40	48	2.8	1579	1	V64590 Tobacco chl DNA.
41	48	2.8	2120	1	X33812 Coding sequence fo
42	48	2.8	1090	1	X37521 Human secreted pro
43	47	2.7	1239	1	N91576 Sequence of clone

44	47	2.7	2549	1	N91467	Sequence of human
45	47	2.7	1834	1	Q04690	Encodes Mammalian
46	47	2.7	372	1	N80489	Cowpea trypsin inh
47	47	2.7	7722	1	Q12023	Norwalk virus geno
48	47	2.7	86	1	Q11760	Self-complementary
49	47	2.7	1793	1	Q15139	Clone pLAC6 encod
50	47	2.7	1410	1	Q15022	Hyoxyamine 6 beta
51	47	2.7	1118	1	Q02063	Deg-1 gene. DNA se
52	47	2.7	1239	1	Q23000	Sequence encoding
53	47	2.7	766	1	N40162	Sequence of prepro
54	47	2.7	1396	1	Q24371	Natural killer cel
55	47	2.7	1078	1	Q25785	Protein synthesis
56	47	2.7	2116	1	Q35987	Tomato hsp80 CDNA
57	47	2.7	1239	1	Q43439	LOI PI CDNA clone
58	47	2.7	222	1	Q46071	Sequence downstre
59	47	2.7	1453	1	Q50573	Asparaginylendopep
60	47	2.7	1640	1	Q50575	Asparaginylendopep
61	47	2.7	1910	1	Q50579	Asparaginylendopep
62	47	2.7	120	1	Q52429	Human subtelomeric
63	47	2.7	1738	1	Q56241	Sequence encoding
64	47	2.7	7753	1	Q56826	Norwalk virus stia
65	47	2.7	1172	1	Q58669	Protein allergen o
66	47	2.7	1393	1	Q66241	Alpha-galactosidas
67	47	2.7	600	1	Q72739	T. ovis vaccine ca
68	47	2.7	1393	1	Q77861	Human alpha-galact
69	47	2.7	458	1	Q90525	Rat Sirt 15 kDa su
70	47	2.7	1748	1	Q88327	Arabidopsis thalia
71	47	2.7	120	1	Q87011	Subtelomeric CDNA
72	47	2.7	399	1	T13044	Cotton fibre-speci
73	47	2.7	767	1	T15821	TBE virus strain N
74	47	2.7	1141	1	T15820	TBE virus strain N
75	47	2.7	600	1	T17934	Taenia ovis anti
76	47	2.7	1582	1	T16831	Human survival mot
77	47	2.7	1560	1	T18828	Human survival mot
78	47	2.7	1558	1	T18255	Human survival mot
79	47	2.7	1582	1	T28259	Survival motor neu
80	47	2.7	399	1	T30261	Cotton fibre cell-
81	47	2.7	1141	1	T44469	Tick-borne enceph
82	47	2.7	767	1	T44470	Tick-borne enceph
83	47	2.7	1364	1	T48099	Human interleukin-
84	47	2.7	3524	1	T46186	CDNA encoding cell
85	47	2.7	1174	1	T39050	Cotton fibre speci
86	47	2.7	399	1	T62620	CDNA encoding anti
87	47	2.7	916	1	T65651	Rat RNA polymerase
88	47	2.7	458	1	T55993	PTH-like peptide D
89	47	2.7	2123	1	T59701	Human origin of re
90	47	2.7	3214	1	T62360	Human origin of re
91	47	2.7	4391	1	T60666	Mouse CD100 anti
92	47	2.7	2589	1	T61590	Human c-IAP1. Nuci
93	47	2.7	399	1	T70051	Cotton fibre speci
94	47	2.7	1868	1	T73502	Poplar 1-aminocycl
95	47	2.7	3214	1	T73287	Human origin of re
96	47	2.7	966	1	T87430	Clone H174-10. Pol
97	47	2.7	120	1	T86338	Subtelomeric CDNA
98	47	2.7	375	1	T84939	Human prostate pro
99	47	2.7	69	1	V02148	Human secreted pro
100	47	2.7	101	1	V00420	3' fragment of clo
101	47	2.7	99	1	T91300	Human M97-2 secret
102	47	2.7	1316	1	T98543	Human chemokine re
103	47	2.7	69	1	T88081	3' portion of CDNA
104	47	2.7	2239	1	V01527	Wheat soluble stia
105	47	2.7	85	1	V05720	Nucleotide sequenc
106	47	2.7	1604	1	V04252	Wild tomato acyltr
107	47	2.7	2230	1	V15704	Tomato ACC synthas
108	47	2.7	1364	1	V15228	Heterodimeric huma
109	47	2.7	144	1	V21239	Human sapiens clone
110	47	2.7	1700	1	V11858	Mus musculus Tub I
111	47	2.7	79	1	V37197	Oligonucleotide se
112	47	2.7	80	1	V26706	Human novel secret
113	47	2.7	1174	1	V39096	Monocomponent endo
114	47	2.7	4358	1	V42653	Nucleotide sequenc
115	47	2.7	79	1	V23414	Human sapiens clone
116	47	2.7	1364	1	V42536	Human interleukin-

117	47	2.7	568	1	V53256	ETH precursor anti
118	47	2.7	2790	1	V32998	Human concentrativ
119	47	2.7	1508	1	V07665	Maize Bx1 cDNA (DI
120	47	2.7	1694	1	V33190	Secreted protein B
121	47	2.7	1753	1	V59556	Human secreted pro
122	47	2.7	2682	1	V63195	CDNA from clone 19
123	47	2.7	916	1	V62000	R. prolixus NO-r p
124	47	2.7	1406	1	V61486	Human secreted pro
125	47	2.7	1114	1	V34315	Human secreted pro
126	47	2.7	1123	1	V34290	Human secreted pro
127	47	2.7	1147	1	V69620	Human secreted pro
128	47	2.7	568	1	V34054	Antisense full len
129	47	2.7	98	1	X00174	Porcine reproducti
130	47	2.7	1376	1	X04325	Human secreted pro
131	47	2.7	1240	1	X06787	Human adult trache
132	47	2.7	1230	1	X07565	Homo sapiens fetal
133	47	2.7	826	1	X27402	Human secreted pro
134	47	2.7	1223	1	X25130	Soybean isoflavone
135	47	2.7	997	1	X52274	Protein PRO244 CDN
136	47	2.7	546	1	X37513	CDNA sequence enco
137	47	2.6	2676	1	002819	Papaya ringspot vi
138	46	2.6	1124	1	003669	Sequence of plasmid
139	46	2.6	3926	1	003736	Plasmid pfghl enco
140	46	2.6	3820	1	003517	Cucumber mosaic vi
141	46	2.6	1423	1	N90249	DNA sequence of ri
142	46	2.6	1923	1	N91039	Human pro-urokinas
143	46	2.6	2427	1	004107	Aequorin gene Bios
144	46	2.6	957	1	004441	Entire porcine tra
145	46	2.6	2671	1	003303	Clone lmd2 encodin
146	46	2.6	3504	1	005304	Sequence of new pl
147	46	2.6	3850	1	N81634	pmQ40 aequorin ge
148	46	2.6	959	1	N81534	Sequence encoding
149	46	2.6	3784	1	005336	Plasmid encoding
150	46	2.6	6020	1	006648	Sequence of CDNA 1
151	46	2.6	3905	1	N70461	Sequence of CDNA 1
152	46	2.6	1801	1	N70524	DNA encoding human
153	46	2.6	1922	1	N70525	Encodes Xenopus Bo
154	46	2.6	2035	1	N70687	Plasmid pMG4B12 us
155	46	2.6	1538	1	010836	Self-complementary
156	46	2.6	748	1	010376	Human plasmidogen
157	46	2.6	90	1	011762	Fragment D of urac
158	46	2.6	6010	1	011998	Human pro-growth h
159	46	2.6	1013	1	012431	Plasmid sequence e
160	46	2.6	3557	1	N60801	Human pre-prolacti
161	46	2.6	3547	1	N60846	Rhizopus PGK1 prom
162	46	2.6	3628	1	N60848	Human pro-urokinas
163	46	2.6	3621	1	N60847	Sequence of a modif
164	46	2.6	2051	1	Q14944	Human Factor XIII
165	46	2.6	2377	1	Q20360	T23 cDNA. Stamen-s
166	46	2.6	941	1	N30062	CA455. Anthr-spec
167	46	2.6	3905	1	025893	Korean hepatitis C
168	46	2.6	347	1	Q27482	NANB virus strain
169	46	2.6	796	1	Q27948	Rape acyl-ACP thio
170	46	2.6	9472	1	Q33282	Human anti-HBs 11g
171	46	2.6	9589	1	Q38218	Sequence of the 3'
172	46	2.6	1710	1	Q47966	Human anti-HBs 11g
173	46	2.6	1066	1	Q49943	Sequence of the 3'
174	46	2.6	289	1	Q53450	T84.12 light chain
175	46	2.6	1041	1	Q54651	T84.12 L4-12-1 11g
176	46	2.6	1041	1	Q54653	Pig TGF-beta-3. Nu
177	46	2.6	2669	1	Q56925	Human SOD. Recombi
178	46	2.6	667	1	Q61566	Human SOD. Recombi
179	46	2.6	688	1	Q61572	Hybrid human/pig S
180	46	2.6	667	1	Q61575	Sequence encoding
181	46	2.6	5994	1	Q65674	Protein kinase (CK
182	46	2.6	2405	1	Q70827	Rat NDF clone 22 D
183	46	2.6	3344	1	Q80228	Human PRNDF-alpha
184	46	2.6	2335	1	Q80216	Human NDF-alpha2b
185	46	2.6	1651	1	Q80218	T2 DNA-polymerase
186	46	2.6	2760	1	Q85425	T4 DNA-polymerase
187	46	2.6	2760	1	Q85426	Flower style-speci
188	46	2.6	826	1	Q79736	Flower style-speci
189	46	2.6	826	1	Q79736	Flower style-speci
190	46	2.6	779	1	Q79737	Flower style-speci
191	46	2.6	2430	1	Q83161	DNA encoding malic
192	46	2.6	1813	1	Q88760	Human ubiquitous n
193	46	2.6	1536	1	Q94111	mML genomic DNA. T
194	46	2.6	2405	1	Q92962	Human HRR25-like c
195	46	2.6	1461	1	T01471	Mammalian mast cel
196	46	2.6	141	1	T06012	Immunodominant tra
197	46	2.6	570	1	T13041	Cotton fibre-speci
198	46	2.6	1641	1	T13041	Human interleukin-
199	46	2.6	1534	1	T15758	OR-1 orphan recept
200	46	2.6	3437	1	T15729	Spinach debranchin
201	46	2.6	10136	1	T31478	Kinetochores protei
202	46	2.6	1837	1	T33458	Rat vas deferens p
203	46	2.6	1657	1	T33852	Human p57 coding s
204	46	2.6	191	1	T13475	Capture probe for
205	46	2.6	570	1	T30258	Cotton fibre cell-
206	46	2.6	3550	1	T18200	Infectious bursal
207	46	2.6	1154	1	T39048	CDNA encoding cell
208	46	2.6	826	1	T64553	Tomato S-ribonucle
209	46	2.6	609	1	T64554	Tomato S-ribonucle
210	46	2.6	903	1	T62617	Cotton fibre speci
211	46	2.6	1166	1	T60872	Cotton fibre speci
212	46	2.6	10288	1	T71522	Human g protein ga
213	46	2.6	2400	1	T49364	Mouse lipopolysacc
214	46	2.6	1113	1	T75765	Cotton fibre speci
215	46	2.6	13414	1	T71321	Full length 3' UTR
216	46	2.6	240	1	T76782	Plasmid pC851 enco
217	46	2.6	140	1	T76781	DNA encoding human
218	46	2.6	1001	1	T73851	Yellowtail tuna DN
219	46	2.6	1426	1	T72272	Mouse receptor ME2
220	46	2.6	1898	1	T79634	Mouse receptor ME2
221	46	2.6	3851	1	T79857	Human NB phox 3' u
222	46	2.6	10195	1	T85320	Alzheimer's diseas
223	46	2.6	6791	1	T85319	Alzheimer's diseas
224	46	2.6	1769	1	T85973	Alzheimer's diseas
225	46	2.6	1420	1	T72167	Alzheimer's diseas
226	46	2.6	1858	1	T72175	Alzheimer's diseas
227	46	2.6	3527	1	T72171	Human immunodefici
228	46	2.6	1505	1	T72177	Malassezia fungus
229	46	2.6	5173	1	T89783	Malassezia fungus
230	46	2.6	728	1	T88876	Malassezia fungus
231	46	2.6	812	1	T88877	Tobacco calcium/ca
232	46	2.6	1776	1	T91744	Al seed rape cyst
233	46	2.6	1577	1	T90169	Intron 8 of human
234	46	2.6	4586	1	T96838	3' fragment of clo
235	46	2.6	259	1	V00423	Cell membrane prot
236	46	2.6	615	1	V02308	Human HMGI-C abert
237	46	2.6	3451	1	V02308	3' nucleotide sequ
238	46	2.6	239	1	V02881	Human presenilin 1
239	46	2.6	241	1	V09114	Plant CCR enzyme D
240	46	2.6	292	1	V21153	Plant CCR enzyme D
241	46	2.6	2205	1	V23878	Human IP-10/Mig re
242	46	2.6	741	1	V23878	Human BHL1 coding
243	46	2.6	740	1	V23878	Human secreted pro
244	46	2.6	1670	1	V26557	Human secreted pro
245	46	2.6	1844	1	V24018	Human secreted pro
246	46	2.6	4931	1	V29584	Human receptor for
247	46	2.6	1711	1	V36317	C. felis esterase,
248	46	2.6	1728	1	V38990	Novel haemopoietic
249	46	2.6	2836	1	V40744	Maize starch-asso
250	46	2.6	2836	1	V40744	Seq ID #7 from DEL
251	46	2.6	2836	1	V40744	Human secreted pro
252	46	2.6	2836	1	V40744	Human secreted pro
253	46	2.6	2836	1	V40744	Human secreted pro
254	46	2.6	2836	1	V40744	Human secreted pro
255	46	2.6	2836	1	V40744	Human secreted pro
256	46	2.6	2836	1	V40744	Human secreted pro
257	46	2.6	2836	1	V40744	Human secreted pro
258	46	2.6	2836	1	V40744	Human secreted pro
259	46	2.6	2836	1	V40744	Human secreted pro
260	46	2.6	2836	1	V40744	Human secreted pro
261	46	2.6	2836	1	V40744	Human secreted pro
262	46	2.6	2836	1	V40744	Human secreted pro

263	46	2	6	2892	1	V58361	Coding sequence fo
264	46	2	6	2280	1	V52604	Human metalloprote
265	46	2	6	2639	1	V52934	Pig transforming g
266	46	2	6	1836	1	V33461	Human T1-receptor
267	46	2	6	3899	1	V63192	cDNA from clone dt
268	45	2	6	577	1	V59806	Human secreted pro
269	46	2	6	1378	1	V59706	Human secreted pro
270	46	2	6	1509	1	V59679	Human secreted pro
271	46	2	6	1196	1	V59663	Human secreted pro
272	46	2	6	1435	1	V59619	Human secreted pro
273	46	2	6	975	1	V61499	Bak binding protei
274	46	2	6	1029	1	V45444	Human chemokine zs
275	46	2	6	1599	1	V34294	Human secreted pro
C 276	46	2	6	100	1	V68827	DNA molecule encod
277	46	2	6	4051	1	V62572	Tumour necrosis fa
278	46	2	6	3964	1	V62673	Tumour necrosis fa
C 279	46	2	6	340	1	V89136	EST clone BR309. N
280	46	2	6	2271	1	V84632	Human secreted pro
281	46	2	6	2276	1	V84583	Human secreted pro
282	46	2	6	2218	1	V84503	Human secreted pro
283	46	2	6	2483	1	V84468	Human secreted pro
284	46	2	6	1686	1	V73012	Human adult brain
285	46	2	6	2294	1	V79584	Rat organic anion
286	46	2	6	1733	1	V81394	Human tumour anti
287	46	2	6	1015	1	X00713	Human secreted pro
288	46	2	6	1020	1	X00681	Human secreted pro
289	46	2	6	1215	1	X00620	Human secreted pro
290	46	2	6	1153	1	V55748	Human secreted pro
291	46	2	6	3213	1	V55742	Human secreted pro
292	46	2	6	645	1	X04348	Human secreted pro
293	46	2	6	1261	1	X04382	Human secreted pro
294	46	2	6	890	1	X04376	Human secreted pro
295	46	2	6	722	1	X06786	Human adult uterus
296	46	2	6	3059	1	V99912	Human salivoadhesin
297	46	2	6	3099	1	V99911	Human salivoadhesin
298	46	2	6	1146	1	X07000	Xenopus Frazzled p
299	46	2	6	2264	1	X20496	Human secreted pro
300	46	2	6	2265	1	X20435	Human secreted pro
301	46	2	6	602	1	X20419	Human secreted pro
302	46	2	6	2888	1	X22123	Human secreted pro
303	46	2	6	1332	1	X02709	Human transaldolas
304	46	2	6	832	1	X30563	DNA encoding a hum
305	46	2	6	1230	1	X22276	Human secreted pro
306	46	2	6	5267	1	X21357	Human BAI3 gene. N
307	46	2	6	1564	1	X22331	Human secreted pro
308	46	2	6	1654	1	X07431	Homo sapiens secre
309	46	2	6	1835	1	X27245	Human CLAR1 coding
310	46	2	6	3375	1	X28358	Human Stat6 coding
311	46	2	6	885	1	X27366	Human secreted pro
312	46	2	6	866	1	X51754	DNA encoding a hum
313	46	2	6	1361	1	X51714	DNA encoding a hum
314	46	2	6	487	1	X30156	Human secreted pro
315	46	2	6	935	1	X51702	DNA encoding a hum
316	46	2	6	1367	1	X51727	DNA encoding a hum
317	46	2	6	1349	1	X24921	Guman ras carboxy-
318	46	2	6	650	1	X37406	Human secreted pro
319	46	2	6	824	1	X37387	Human secreted pro
320	46	2	6	773	1	X37388	Human secreted pro
321	46	2	6	604	1	X37373	Human secreted pro
322	46	2	6	2487	1	X33813	Coding sequence fo
323	46	2	6	7797	1	X33180	Cowpox virus bsr f
324	46	2	6	6644	1	X33181	Base sequence of t
325	46	2	6	7372	1	X33182	Base sequence of t
326	46	2	6	7996	1	X33184	Base sequence of t
327	46	2	6	980	1	X37465	Human secreted pro
328	46	2	6	971	1	X37469	Human secreted pro
329	46	2	6	1510	1	X03740	Human myb related
330	46	2	6	659	1	N90103	Human prealbumin c
331	45	2	6	4180	1	N91773	Rat androgen recep
332	45	2	6	1973	1	N91045	Gene encoding bov1
333	45	2	6	1493	1	O04492	Sequence encoding
334	45	2	6	960	1	O03293	Recombinant DNA en
335	45	2	6	2400	1	O05055	Placenta-specific

Sequence of cDNA e
CBT-20 cDNA encodi
Placenta-specific
p150.95 alpha subu
PMX34 cDNA insert.
GTP-gene from Chin
MHC class II antiq
H77 membrane prote
Shuttle vector PMU
cDNA sequence enco
Sequence encoding
Plasmid pMG3C9 use
Self-complementary
Caldesmon-like pol
Shuttle vector pSW
Plasmid pAU157 ins
Sequence of plasmid
Sequence of the br
plasmid pAU157 seq
Sequence encoding
Heparin-binding gr
NF-YB. Expression
Sequence of the S3
Sequence encoding
Clone W264 encodin
Clone W282 encodin
pSW6 expression ve
Alpha-GalNAc from
pSP6-SFV4 RNA tran
Tyrosine Kinase re
Glia activating fa
Human kidney PTH/p
Cyclin D1 gene. Re
Human heregulin-be
ppcdHP23 coding se
pVMO31. DNA sequen
Tomato acid invert
Sequence of part o
pSW6 for expressio
Recombinant human
Human anti-HBs hea
Human cyclin D1. N
Sequence encoding
Carcinomaembryonic a
Tomato acid invert
poly(dA) 50mer prob
Wheat leaf ADP-glu
Apo-B RNA editing
Human basigin I im
Human c-myc far up
Human c-myc far up
Hepatitis C virus
Gibberellin-20-ox
RP-8 progammmed ce
Murine glycosylati
Hepatitis C virus
Spinoecerebellar at
Sequence encoding
TIMP-3 metalloprot
Sequence encoding
DNA encoding phosph
Mylon aconitase cd
Tomato acid invert
Achtalia fulica cd
Human kidney PTH/p
Human kidney PTH/p
Integrin subunit b
Cancer suppressor
Human interleukin-
Human DRADA gene 6
Cystathionine gamma
Coding sequence fo
Human milk kappa-c
Human L-asparagina

409	45	2.6	1623	1	T14925	T cell replacing f
410	45	2.6	2158	1	T15228	Tumor necrosis fac
411	45	2.6	1121	1	T8129	Nuclear envelope-a
412	45	2.6	1425	1	T35220	Cytoplasmic antipr
413	45	2.6	1016	1	T29401	Lipase coding sequ
414	45	2.6	141	1	T34595	Probe for detecti
415	45	2.6	635	1	T38390	Murine glycosylati
416	45	2.6	1602	1	T45865	CDNA encoding C. h
417	45	2.6	1630	1	T43204	Human mitogen-acti
418	45	2.6	1964	1	T46064	Carcinohembryonic a
419	45	2.6	3736	1	T45937	Human hypoxia indu
420	45	2.6	940	1	T58281	Arabidopsis Sfz po
421	45	2.6	1276	1	T42071	Human nuclear prol
422	45	2.6	3040	1	T42853	Phospholipase D pr
423	45	2.6	780	1	T59912	Human transcriptio
424	45	2.6	2001	1	T59700	pH-like peptide D
425	45	2.6	3871	1	T65952	Human Ob receptor
426	45	2.6	1849	1	T74042	Soybean thiol prot
427	45	2.6	90	1	T47082	Synthetic DNA prob
428	45	2.6	90	1	T47081	Synthetic DNA prob
429	45	2.6	5894	1	T85474	hABC3 cDNA sequenc
430	45	2.6	1713	1	T85993	Maize 5-enolpyruv
431	45	2.6	1713	1	T86000	Maize 5-enolpyruv
432	45	2.6	558	1	T64563	Mouse thymus and a
433	45	2.6	204	1	T61968	Detection probe fo
434	45	2.6	249	1	T89608	Hepatitis C virus
435	45	2.6	227	1	T89600	Hepatitis C virus
436	45	2.6	270	1	T89602	Hepatitis C virus
437	45	2.6	260	1	T89501	Hepatitis C virus
438	45	2.6	356	1	T89597	Hepatitis C virus
439	45	2.6	257	1	T89599	Hepatitis C virus
440	45	2.6	1023	1	T70132	Max-interacting pr
441	45	2.6	263	1	T85891	Malassezia fungus
442	45	2.6	1512	1	T72172	Alzheimer's disease
443	45	2.6	882	1	T72173	Alzheimer's disease
444	45	2.6	6671	1	T89799	Human double stran
445	45	2.6	3040	1	T85509	Phospholipase D en
446	45	2.6	2158	1	T94651	TNF-RI-D ligand p
447	45	2.6	2628	1	T01060	Human PKR gene. Sc
448	45	2.6	2481	1	T87940	Rat neuropeptide y
449	45	2.6	2604	1	T73595	Rat neuropeptide y
450	45	2.6	1040	1	V02818	Human RHAM genom
451	45	2.6	3114	1	V02800	Human RHAM cDNA.
452	45	2.6	1441	1	T90173	Oil seed rape cyst
453	45	2.6	106	1	V00430	3' fragment of clo
454	45	2.6	720	1	V00421	Clone B219-2. Isol
455	45	2.6	863	1	T91302	Human H1075-1 secr
456	45	2.6	877	1	V04635	Human cytoplasmic
457	45	2.6	5894	1	V16307	CDNA encoding huma
458	45	2.6	1817	1	V04064	CDNA for wild type
459	45	2.6	2096	1	V15795	Human retinoid rec
460	45	2.6	2233	1	V10120	CDNA encoding full
461	45	2.6	6535	1	V16345	3' nucleotide sequ
462	45	2.6	208	1	V09116	Rat GAIIR2 receptor
463	45	2.6	1958	1	V10328	Rat FRAG1 coding s
464	45	2.6	1780	1	V04269	Human cycloin DI on
465	45	2.6	1335	1	V20473	Secreted protein C
466	45	2.6	84	1	V04275	Nucleotide sequen
467	45	2.6	84	1	V09269	Pea plastidial pro
468	45	2.6	2182	1	V00731	Chimeric protease
469	45	2.6	1136	1	V12373	Human haematopoiet
470	45	2.6	1136	1	V20812	Equine arteritis v
471	45	2.6	12827	1	V09036	Equine arteritis v
472	45	2.6	15528	1	V09039	Equine arteritis v
473	45	2.6	889	1	V19304	Human p26 cDNA seq
474	45	2.6	69	1	V21236	Homo sapiens clone
475	45	2.6	117	1	V21232	Homo sapiens clone
476	45	2.6	69	1	V21226	Homo sapiens clone
477	45	2.6	86	1	V11441	Human secreted pro
478	45	2.6	1454	1	V23913	Plant CAD enzyme D
479	45	2.6	1474	1	V23914	Plant CAD enzyme D
480	45	2.6	1749	1	V21800	Human spleen-deriv
481	45	2.6	2065	1	V29062	BRCA1 modulator pr

555	45	2.6	1717	1	V08170	MMP19 coding seque
556	45	2.6	1152	1	V62754	Human secreted pro
557	45	2.6	378	1	V90045	EST clone CW150.
558	45	2.6	283	1	V89992	EST clone CW762. N
559	45	2.6	2010	1	V08391	Human Parathyroid
560	45	2.6	579	1	V82799	ATG-1100 (allotraf
561	45	2.6	1181	1	V81616	Cucumber peroxidase
562	45	2.6	2017	1	V08830	Gene No. 20 encodi
563	45	2.6	1705	1	V08856	Human secreted huma
564	45	2.6	1705	1	V84516	Human secreted pro
565	45	2.6	941	1	V84586	Human secreted pro
566	45	2.6	1358	1	V84593	Human secreted pro
567	45	2.6	1637	1	V84578	Human secreted pro
568	45	2.6	472	1	V84530	Human secreted pro
569	45	2.6	1059	1	V84474	Human secreted pro
570	45	2.6	588	1	V84464	Human secreted pro
571	45	2.6	1210	1	V84425	Human secreted pro
572	45	2.6	467	1	V64423	Human secreted pro
573	45	2.6	77	1	V77194	Mouse developing 1
574	45	2.6	1521	1	V73000	Staphylococcus aur
575	45	2.6	1908	1	V73002	Human foetal brain
576	45	2.6	3076	1	V73003	Human adult brain
577	45	2.6	2171	1	V73006	Human adult brain
578	45	2.6	1181	1	V62867	Human foetal brain
579	45	2.6	4456	1	V81748	Cucumber peroxidase
580	45	2.6	1425	1	X00679	Human ALP encoding
581	45	2.6	1668	1	X00664	Human secreted pro
582	45	2.6	1507	1	X00656	Human secreted pro
583	45	2.6	2788	1	V55745	Human secreted pro
584	45	2.6	1813	1	V70895	Human secreted pro
585	45	2.6	1125	1	V80740	CDNA encoding an a
586	45	2.6	1337	1	X04345	Human secreted pro
587	45	2.6	667	1	X04334	Human secreted pro
588	45	2.6	2541	1	X04343	Human secreted pro
589	45	2.6	1322	1	X04402	Human secreted pro
590	45	2.6	2298	1	X06781	Human testis secre
591	45	2.6	1180	1	V84347	Human Apo-2DCR CDN
592	45	2.6	74	1	V99725	Human adult testis
593	45	2.6	1554	1	X16675	Xenopus WA545 prot
594	45	2.6	2880	1	X18193	hNBC3a coding sequ
595	45	2.6	1602	1	X07066	Human mtogen acti
596	45	2.6	3912	1	X18192	hNBC3 coding sequ
597	45	2.6	1761	1	X20412	Human secreted pro
598	45	2.6	1071	1	X20414	Human secreted pro
599	45	2.6	4765	1	X09003	Mouse semaphorin r
600	45	2.6	4780	1	X09004	Mouse semaphorin r
601	45	2.6	687	1	X30410	DNA encoding a hum
602	45	2.6	1095	1	X30366	DNA encoding a hum
603	45	2.6	687	1	X30351	DNA encoding a hum
604	45	2.6	513	1	X22214	Human secreted pro
605	45	2.6	492	1	X22223	Human secreted pro
606	45	2.6	1613	1	X26196	DNA sequence of q1
607	45	2.6	2634	1	X27062	S. tuberculosis isom
608	45	2.6	3684	1	X21560	Human N-arginine d
609	45	2.6	2152	1	X22002	Human MDL endonuc
610	45	2.6	2345	1	X18951	Human PIGR-1 encod
611	45	2.6	337	1	X25860	Clone Pegen-42 seq
612	45	2.6	588	1	X07427	Homo sapiens secre
613	45	2.6	2773	1	X07428	Homo sapiens secre
614	45	2.6	791	1	X19493	Human secreted pro
615	45	2.6	5668	1	X27262	Prostate-tumour de
616	45	2.6	3848	1	X27277	Prostate-tumour de
617	45	2.6	1129	1	X27340	Human secreted pro
618	45	2.6	921	1	X27330	Human secreted pro
619	45	2.6	553	1	X51720	DNA encoding a hum
620	45	2.6	1032	1	X30155	Human secreted pro
621	45	2.6	340	1	X51737	DNA encoding a hum
622	45	2.6	1442	1	X51744	DNA encoding a hum
623	45	2.6	990	1	X51748	DNA encoding a hum
624	45	2.6	978	1	X51732	DNA encoding a hum
625	45	2.6	986	1	X51735	DNA encoding a hum
626	45	2.6	3495	1	X33814	Coding sequence fo
627	45	2.6	1493	1	X33994	Human osteopontin

628	45	2.6	1547	1	X33448	Oryza sativa L. pi
629	45	2.6	1984	1	X52252	Protein PRO271 CDN
630	45	2.6	1503	1	X52253	Protein PRO272 CDN
631	45	2.6	2822	1	X52222	Protein PRO228 CDN
632	45	2.6	1060	1	X37464	Human secreted pro
633	45	2.6	1448	1	X37462	Human secreted pro

ALIGNMENTS

RESULT	1
ID	080521
AC	080521: standard: cDNA; 1737 BP.
DT	18-JUL-1995 (first entry)
DE	Human monocyte Pf4AR cDNA.
KW	Interleukin-8 receptor; IL-8 receptor; Pf4AR;
KW	platelet factor superfamily receptor; monocyte; chemotactic;
KW	inflammation; inflammatory disease; arthritis; emphysema; cystic;
OS	Homo sapiens; colitis; bronchitis; meningitis; therapeutic; ss.
FT	Key Location/Qualifiers
FT	cds 91..1149
FT	/*tag= a
PN	MO9428931-A.
PD	22-DEC-1994.
PF	07-JUN-1994.
PR	11-JUN-1993; US-076093.
PA	(GETH) GENENTECH INC.
PI	Chuncharapal A, Hebert C, Kim KJ, Lee J;
DR	WPI; 95-036114/05.
PT	P-PSDB: R68812.
PT	Treatment of inflammatory disorders - by administering an
PT	antibody capable of binding a platelet factor 4 superfamily
PT	receptor polypeptide
PS	Disclosure; Page 54-56; 83pp; English.
CC	2 Pf4AR members were identified by probing lambda libraries from
CC	CC human monocytic-like cell line HL-60 and human peripheral blood
CC	CC lymphocytes using a large fragment of IL-8 receptor DNA (full
CC	CC sequence given in Q80520). The nucleotide sequences of the 2
CC	CC Pf4ARs are given in Q80521 and Q80522, and their respective
CC	CC amino acid sequences in R68812 and R68813.
SQ	Sequence 1737 BP; 454 A; 411 C; 373 G; 499 T;

Query Match 100.0%; Score 1737; DB 1; Length 1737;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1737; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GAATTCAGTGTCTGCTGCGCGCGCGCAAGTGACGCGAGGCGCTGAGTCTCAGTA	60
DB	1	GAATTCAGTGTCTGCTGCGCGCGCGCGCAAGTGACGCGAGGCGCTGAGTCTCAGTA	60
QY	61	GCCACCGCATCTGAGAACACCGCGGTTACCATGGAGGGATCAGTATATACATTCAGAT	120
DB	61	GCCACCGCATCTGAGAACACCGCGGTTACCATGGAGGGATCAGTATATACATTCAGAT	120
QY	121	AACATACCGAGGAATGAGGCTCAGGAGCATGATGATCCATGAAAGAACCTGTTCCGT	180
DB	121	AACATACCGAGGAATGAGGCTCAGGAGCATGATGATCCATGAAAGAACCTGTTCCGT	180
QY	181	GAAGAAATGCTAATTTCAATTAATTTCTGCGCCACCATCTACTCATCATCTTCTTA	240
DB	181	GAAGAAATGCTAATTTCAATTAATTTCTGCGCCACCATCTACTCATCATCTTCTTA	240
QY	241	ACTGCGATTGTGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	300
DB	241	ACTGCGATTGTGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	300
QY	301	AGCATGACGACAAAGTACAGGCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTG	360
DB	301	AGCATGACGACAAAGTACAGGCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTG	360

```

OY 361 CTTCCTTCTGGGCACTGTGATGCCGTGGCAAACTGCTACTTTGGGAACCTCTCTATGCAAG 420
DB 361 CTTCCTTCTGGGCACTGTGATGCCGTGGCAAACTGCTACTTTGGGAACCTCTCTATGCAAG 420
OY 421 GCAGTCATCTCATCTACAGATCAACCTCTACAGCAGTGTCCATCCATCCGACCTTCATC 480
DB 421 GCAGTCATCTCATCTACAGATCAACCTCTACAGCAGTGTCCATCCATCCGACCTTCATC 480
OY 481 AGCTGAGCAGCTACTGCGCATCTGTCAGCCCAACAAGTCTAGAGGCCAAGAGAGCTG 540
DB 481 AGCTGAGCAGCTACTGCGCATCTGTCAGCCCAACAAGTCTAGAGGCCAAGAGAGCTG 540
OY 541 TTGGCTGAAAAAGTGTCTATGTTGGGCTGTGATCCCTGCCCCCTGCTGCTATTCCTCC 600
DB 541 TTGGCTGAAAAAGTGTCTATGTTGGGCTGTGATCCCTGCCCCCTGCTGCTATTCCTCC 600
OY 601 GACTTCATCTTTGCCAAGCTAGTGAAGGAGATGACAGATATATCTGTGACCGCTTCAC 660
DB 601 GACTTCATCTTTGCCAAGCTAGTGAAGGAGATGACAGATATATCTGTGACCGCTTCAC 660
OY 661 CCCATGACTTGTGGTGTGTGTGTTCCAGTTTCAGCAGATCATGTGTGCGCTTATCCTG 720
DB 661 CCCATGACTTGTGGTGTGTGTGTTCCAGTTTCAGCAGATCATGTGTGCGCTTATCCTG 720
OY 721 CCTGGTATTTGATCTCTCTCTCTATTTGCAATATCATCTCCAGCTGTCACATCCCAAG 780
DB 721 CCTGGTATTTGATCTCTCTCTCTATTTGCAATATCATCTCCAGCTGTCACATCCCAAG 780
OY 781 GGGCACCAGAGCGCAAGCGCCCTCAAGACACAGTCACTCATCTGCTGCTTCTTTCGCC 840
DB 781 GGGCACCAGAGCGCAAGCGCCCTCAAGACACAGTCACTCATCTGCTGCTTCTTTCGCC 840
OY 841 TGTGGCTGCTTACTATCATTTGGATGATCAGATCGATCTCTTCATCTCTCTGGAATCATC 900
DB 841 TGTGGCTGCTTACTATCATTTGGATGATCAGATCGATCTCTTCATCTCTCTGGAATCATC 900
OY 901 AAGCAAGGCTGTGATTTGAGAACACTGTGCAAGTGTGATTTCCATCACCAGGCGCTA 960
DB 901 AAGCAAGGCTGTGATTTGAGAACACTGTGCAAGTGTGATTTCCATCACCAGGCGCTA 960
OY 961 GCTTCTTCTCACTGTGTCTGTAACCCCACTCTATGCTTCTTGGAGCCAAATTTAAA 1020
DB 961 GCTTCTTCTCACTGTGTCTGTAACCCCACTCTATGCTTCTTGGAGCCAAATTTAAA 1020
OY 1021 ACCTTGTCCAGCAGCAGCTACCTCTGTGACAGAGGCTCCAGCTCAAGATTCCTCTCC 1080
DB 1021 ACCTTGTCCAGCAGCAGCTACCTCTGTGACAGAGGCTCCAGCTCAAGATTCCTCTCC 1080
OY 1081 AAAGGAAAGCGAGTGGAGCATTCATCTGTTCCAGTGTGAGTCTCAAGTTTTCAC 1140
DB 1081 AAAGGAAAGCGAGTGGAGCATTCATCTGTTCCAGTGTGAGTCTCAAGTTTTCAC 1140
OY 1141 TCCAGCTTACACAGATGTAAAGACTTTTTTTATACGATTAATTAATTTTTTAAGTT 1200
DB 1141 TCCAGCTTACACAGATGTAAAGACTTTTTTTATACGATTAATTAATTTTTTAAGTT 1200
OY 1201 ACACATTTTTCAGATATAAAGACTACCAATATTTACAGTTTATGCTGTGGAT 1260
DB 1201 ACACATTTTTCAGATATAAAGACTACCAATATTTACAGTTTATGCTGTGGAT 1260
OY 1261 TTTTGTCTGTGTGTTCTTATGTTTGTGAGTTTATGACTTATTAATAATTTTT 1320
DB 1261 TTTTGTCTGTGTGTTCTTATGTTTGTGAGTTTATGACTTATTAATAATTTTT 1320
OY 1321 TTTTGTCTGTGTGTTCTTATGTTTGTGAGTTTATGACTTATTAATAATTTTT 1380
DB 1321 TTTTGTCTGTGTGTTCTTATGTTTGTGAGTTTATGACTTATTAATAATTTTT 1380
OY 1381 GTCTGCTGTAGAGCTGTAGAAAAAGGAAGTGAACATTCAGAGCGGTGTAGTGAATCAG 1440
DB 1381 GTCTGCTGTAGAGCTGTAGAAAAAGGAAGTGAACATTCAGAGCGGTGTAGTGAATCAG 1440

```

```

OY 1441 TAAAGTAGAAATGATCCCGAGCTGTTATGATAGATAATCTCTCCATTCCTCGTGAC 1500
DB 1441 TAAAGTAGAAATGATCCCGAGCTGTTATGATAGATAATCTCTCCATTCCTCGTGAC 1500
OY 1501 GTTTTCTCTGTTCTTAAAGCTGATTTTGTGTAGAGATGAGCAGTATTAACCAAGCCC 1560
DB 1501 GTTTTCTCTGTTCTTAAAGCTGATTTTGTGTAGAGATGAGCAGTATTAACCAAGCCC 1560
OY 1561 AAAGTGTATGAAGATCTGTTTTCAGTTTTCAGAGAGGAGTTGATTTTCAGCACCTAC 1620
DB 1561 AAAGTGTATGAAGATCTGTTTTCAGTTTTCAGAGAGGAGTTGATTTTCAGCACCTAC 1620
OY 1621 AGTGTACGCTTGTATTAAGTTGTATTAAGTATTAAGTATTAAGTATTAAGTATTAAG 1680
DB 1621 AGTGTACGCTTGTATTAAGTTGTATTAAGTATTAAGTATTAAGTATTAAGTATTAAG 1680
OY 1681 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCCGACACAGCTGGAATTC 1737
DB 1681 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCCGACACAGCTGGAATTC 1737

RESULT 2
ID 099007 standard; cDNA; 1737 BP.
AC 099007;
AT 26-MAR-1996 (first entry)
DE Chemokine superfamily receptor coding sequence.
KW Interleukin; IL-8; inflammation; psoriasis; dermatitis;
KW Rheumatoid arthritis; inflammatory bowel disease;
KW Chronic lung inflammation; treatment; antibody;
KW affinity purification; detection; ss.
OS Homo sapiens.
PN US5440021-A.
PD 08-AUG-1995.
PR 29-MAR-1991; US-677211.
PR 25-FEB-1994; US-202056.
PA (CHUN)/ CHUNTHARAPAI A.
PA (HEBE)/ HEBERT C.
PA (KIMK)/ KIM K J.
PA (LEEJ)/ LEE J.
PI Chuntharapai A., Hebert C., Kim KJ, Lee J;
PI WPI: 95-283151/37.
DR P-PSDB: R80757.
PT New antibodies against interleukin 8 type B receptor - used to treat
PT or prevent inflammation, also for detecting receptor expression and
PT purification.
PS Example 2: Columns 47-50: 62bp: English.
CC Antibodies directed against the interleukin-8 receptor B can be used
CC to treat or prevent inflammation e.g. psoriasis, dermatitis,
CC Rheumatoid arthritis and particularly inflammatory bowel disease and
CC chronic lung inflammation. When immobilised, these antibodies may
CC be used to detect interleukin-8 receptor B expression in cells and
CC tissues and for affinity purification of interleukin-8 receptor B
CC from cells. This sequence is an additional chemokine superfamily
CC receptor which was identified by probing lambda libraries of genomic
CC DNA from a human monocyte-like cell line (L-60) and human peripheral
CC blood lymphocytes using a large fragment of the interleukin-8 type
CC A receptor DNA (See 099006).
SO Sequence 1737 BP; 454 A; 411 C; 373 G; 499 T;

Query Match 100.0%; Score 1737; DB 1: Length 1737;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1737; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GAATTCAGTGTGCTGGGCGGCGGCAAGTGAAGCGGAGGCTGAGTCCAGTA 60
DB 1 GAATTCAGTGTGCTGGGCGGCGGCAAGTGAAGCGGAGGCTGAGTCCAGTA 60
OY 61 GCCACCGCATCTGTGAGAACCGAGGTTTACCATGAGAGGATCATATACACTTACAT 120
DB 61 GCCACCGCATCTGTGAGAACCGAGGTTTACCATGAGAGGATCATATACACTTACAT 120

```

OY	121	AAATACCCAGGAAATGGGCTCAGGGGATATAGATCTCATAGAGAAACCCGTTCCGT	180
Db	121	AACATACCCAGGAAATGGGCTCAGGGGATATAGATCTCATAGAGAAACCCGTTCCGT	180
OY	181	GAGAGAAATGCTAATTTCATATAAATCTTCCTGCCACATCTACATCCATCTCTTA	240
Db	181	GAGAGAAATGCTAATTTCATATAAATCTTCCTGCCACATCTACATCCATCTCTTA	240
OY	241	ACTGGCATTTGGGCAATGCAATGGTCAATCTGGTCAATGGGTTACCGAGAAACTGAGA	300
Db	241	ACTGGCATTTGGGCAATGCAATGGTCAATCTGGTCAATGGGTTACCGAGAAACTGAGA	300
OY	301	AGCATGACGGCAAGTACAGGCTGCACCTCTCATGTGGCCGACCTCTCTTTGATCAGG	360
Db	301	AGCATGACGGCAAGTACAGGCTGCACCTCTCATGTGGCCGACCTCTCTTTGATCAGG	360
OY	361	CTTCCCTTCCTGGGCAATGATGGCTGGCCAACTGGTACTTTGGGAATCTCTATGACAG	420
Db	361	CTTCCCTTCCTGGGCAATGATGGCTGGCCAACTGGTACTTTGGGAATCTCTATGACAG	420
OY	421	GCAGTCATGTCATCTACAGATCAACCTTACAGAGTGTCTCATCTGGCTTATC	480
Db	421	GCAGTCATGTCATCTACAGATCAACCTTACAGAGTGTCTCATCTGGCTTATC	480
OY	481	AGTGTGACCGCTACCTGGGCATCTGTCACAGCCACCAAGTACAGAGGCAAGAGTG	540
Db	481	AGTGTGACCGCTACCTGGGCATCTGTCACAGCCACCAAGTACAGAGGCAAGAGTG	540
OY	541	TTGGCTGAAAAGTGGTCTATGTTGGGCTGTGATCCCTGCTCTGTGACTATCC	600
Db	541	TTGGCTGAAAAGTGGTCTATGTTGGGCTGTGATCCCTGCTCTGTGACTATCC	600
OY	601	GACTTCATCTTGGCACAAGTACAGAGGCAAGATGACAGATATGTCGTACCGCTTAC	660
Db	601	GACTTCATCTTGGCACAAGTACAGAGGCAAGATGACAGATATGTCGTACCGCTTAC	660
OY	661	CCCAATGACTTGTGGGTGGTGGTTGTTCCAGATTACAGACATCATGTTGGCTTATCCG	720
Db	661	CCCAATGACTTGTGGGTGGTGGTTGTTCCAGATTACAGACATCATGTTGGCTTATCCG	720
OY	721	CTGTGATGTGCATACCTGCTCTCTCTATGTCATATCATCTCCAAGTGTGCACATCCAG	780
Db	721	CTGTGATGTGCATACCTGCTCTCTCTATGTCATATCATCTCCAAGTGTGCACATCCAG	780
OY	781	GGCCACCAAGGCGCAGAGGCGCTCAGACACAGTATCTCATCTGTGGCTTCTGCC	840
Db	781	GGCCACCAAGGCGCAGAGGCGCTCAGACACAGTATCTCATCTGTGGCTTCTGCC	840
OY	841	TGTTGGCTGCCCTACTACATTTGGGATCAGCATGACTCCTTATCCTCTGGAAATCATC	900
Db	841	TGTTGGCTGCCCTACTACATTTGGGATCAGCATGACTCCTTATCCTCTGGAAATCATC	900
OY	901	AAGCAAGGGGTGAGTTTGGAACACTGTGCAACATGATTTTCATACCGAGGCCCTA	960
Db	901	AAGCAAGGGGTGAGTTTGGAACACTGTGCAACATGATTTTCATACCGAGGCCCTA	960
OY	961	GCCTTCTTCCACTGTTGTGTGAACCCCATCTGTATGCTTTCTCTGAGCCAAATTTAA	1020
Db	961	GCCTTCTTCCACTGTTGTGTGAACCCCATCTGTATGCTTTCTCTGAGCCAAATTTAA	1020
OY	1021	ACCTCTGCCAGACGCACTCACTCTGTGAGAGAGAGGTTCCAGGCTCAAGTCTCTCC	1080
Db	1021	ACCTCTGCCAGACGCACTCACTCTGTGAGAGAGAGGTTCCAGGCTCAAGTCTCTCC	1080
OY	1081	AAAGCAAGCGAGGTGAGCATATCATCTGTTCCAGTGAATCTGAGATCTTCAAGTTTAC	1140
Db	1081	AAAGCAAGCGAGGTGAGCATATCATCTGTTCCAGTGAATCTGAGATCTTCAAGTTTAC	1140
OY	1141	TCACGCTAACACAGATGTAAGACATTTTTTTATACGATTAATAACTTTTTTAAGTT	1200
Db	1141	TCACGCTAACACAGATGTAAGACATTTTTTTATACGATTAATAACTTTTTTAAGTT	1200

ID	Accession	Gene	Location/Qualifiers	Length
QY	1201	ACACATTTTCAGATATAAAGACACCAATATTTGACAGTTTATTGCTTTGGAT		1280
Db	1201	ACACATTTTCAGATATAAAGACACCAATATTTGACAGTTTATTGCTTTGGAT		1260
QY	1261	TTTGTCTGTGTTCTTTAGTTTGTGCAAGTTTAATGACCTTAATTATATAAATTTT		1320
Db	1261	TTTGTCTGTGTTCTTTAGTTTGTGCAAGTTTAATGACCTTAATTATATAAATTTT		1320
QY	1321	TTTGTTCATATTGATGTGTGCTTAGGACAGACCTGTGCCAAGTCTTAGTGTCTGAT		1380
Db	1321	TTTGTTCATATTGATGTGTGCTTAGGACAGACCTGTGCCAAGTCTTAGTGTCTGAT		1380
QY	1381	GTCCTCGTAGTAGACCTGTAAAGAAAGGAACCTGAACATTCACAAGGCTGTAGTGAATCAG		1440
Db	1381	GTCCTCGTAGTAGACCTGTAAAGAAAGGAACCTGAACATTCACAAGGCTGTAGTGAATCAG		1440
QY	1441	TAAAGCTAGAAATGATCCCGACGCTGTTATGCAATGATATCTTCATTCCTCCGTGAAAC		1500
Db	1441	TAAAGCTAGAAATGATCCCGACGCTGTTATGCAATGATATCTTCATTCCTCCGTGAAAC		1500
QY	1501	GTTTTCCTGTTCTTAAAGAGTATTTGCTGTAGAAATGACACTTAAACCAAGGCC		1560
Db	1501	GTTTTCCTGTTCTTAAAGAGTATTTGCTGTAGAAATGACACTTAAACCAAGGCC		1560
QY	1561	AAATGTGTATAGAAATGCTGTTTTCAGTTTTCAGAGATGGGTTGATTTTCAGCACTAC		1620
Db	1561	AAATGTGTATAGAAATGCTGTTTTCAGTTTTCAGAGATGGGTTGATTTTCAGCACTAC		1620
QY	1621	AGTGTACAGCTCTGTATTAAAGTGTGTAATTAAGATTCATGTTAAACTTTAAAAA		1680
Db	1621	AGTGTACAGCTCTGTATTAAAGTGTGTAATTAAGATTCATGTTAAACTTTAAAAA		1680
QY	1681	AAAAAAGTCTGTATTAAAGTGTGTAATTAAGATTCATGTTAAACTTTAAAAA		1737
Db	1681	AAAAAAGTCTGTATTAAAGTGTGTAATTAAGATTCATGTTAAACTTTAAAAA		1737
RESULT	3			
Q29506				
AC	Q29506			
DC	12-MAR-1993	(first entry)		
DE	New platelet factor 4 receptor superfamily member PF4ARI.			
KM	IL-8R: G-protein coupled receptor family; rhodopsin superfamily;			
OS	Homo sapiens.			
FT	Key	Location/Qualifiers		
FT	cds	91..1149		
FT	/*tag=	a		
FT	/Product=	PF4ARI		
PN	W09217497-A.			
PD	15-OCT-1992.			
PE	23-MAR-1992:	U02317.		
PR	29-MAR-1991:	US-677211.		
PR	19-DEC-1991:	US-810782.		
PA	(GETH) GENENTECH INC.			
PI	Holmes WE, Lee J, Wood WT;			
DR	WPI: 92-366191/44.			
DR	P-PSDB: R27792.			
PT	Isolated human platelet factor 4 super-family receptor			
PT	polypeptide and corresp. antibodies and DNA - useful as			
PT	diagnostic and screening agents, and for treating inflammation or			
PT	PF4AR-mediated disorders			
PS	Claim 7: Fig 4: 78pp: English.			
CC	The IL-8 receptor cDNA sequence was isolated (see Q29505) and a			
CC	874bp sub-fragment of the coding sequence was used as a probe to			
CC	screen human cell line HL60 and human peripheral blood lymphocyte			
CC	cDNA libraries. Two new gene sequences were found that are clearly			
CC	related to the IL-8 receptor. One of these was contained in			
CC	combined clone 8tr.20.15 and is predicted to encode an amino acid			
CC	sequence which is 34% identical with both the high and low affinity			
CC	IL-8 receptors. See also Q31107.			
CC	Sequence 1737 BP: 457 A: 412 C: 370 G: 498 T:			

Query Match 88.3%; Score 1533; DB 1; Length 1737;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1733; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GAATTCAGTGTCTGGCGCGCGGCAAAAGTACGCCGAGGAGGCGTGTCTCCAGTA 60
 1 GAATTCAGTGTCTGGCGCGCGGCAAAAGTACGCCGAGGAGGCGTGTCTCCAGTA 60
 DB 1 GAATTCAGTGTCTGGCGCGCGGCAAAAGTACGCCGAGGAGGCGTGTCTCCAGTA 60
 QY 61 GCCACCGCATCTGGAGAACGCGGTACCATGAGGGGATCAGTATATACATTCAGAT 120
 61 GCCACCGCATCTGGAGAACGCGGTACCATGAGGGGATCAGTATATACATTCAGAT 120
 DB 61 GCCACCGCATCTGGAGAACGCGGTACCATGAGGGGATCAGTATATACATTCAGAT 120
 QY 121 AACTACACCGAGAAATGGGCTCAGGGGACTATGACTCCATGAAGAAACCTGTTCCGT 180
 121 AACTACACCGAGAAATGGGCTCAGGGGACTATGACTCCATGAAGAAACCTGTTCCGT 180
 DB 121 AACTACACCGAGAAATGGGCTCAGGGGACTATGACTCCATGAAGAAACCTGTTCCGT 180
 QY 131 GAGAAATATGCTATTTCAATATAAATCTTCTGCGCCACATCTACTCATCATCTTTCTTA 240
 131 GAGAAATATGCTATTTCAATATAAATCTTCTGCGCCACATCTACTCATCATCTTTCTTA 240
 DB 131 GAGAAATATGCTATTTCAATATAAATCTTCTGCGCCACATCTACTCATCATCTTTCTTA 240
 QY 241 ACTGCAATGTGGGCAATGATGTCATCTGTGTCATGGGTTACGAAAGAACTGAGA 300
 241 ACTGCAATGTGGGCAATGATGTCATCTGTGTCATGGGTTACGAAAGAACTGAGA 300
 DB 241 ACTGCAATGTGGGCAATGATGTCATCTGTGTCATGGGTTACGAAAGAACTGAGA 300
 QY 301 AGCATGACGACAGTACAGGCTGCACCTGTCAGTGGCGGACCTCTTTGTATCAGC 360
 301 AGCATGACGACAGTACAGGCTGCACCTGTCAGTGGCGGACCTCTTTGTATCAGC 360
 DB 301 AGCATGACGACAGTACAGGCTGCACCTGTCAGTGGCGGACCTCTTTGTATCAGC 360
 QY 361 CTTCCTCTTGGGAGTGTGATGCGTGGCAAACTGTACTTTGGGAACTTCCATGCAAG 420
 361 CTTCCTCTTGGGAGTGTGATGCGTGGCAAACTGTACTTTGGGAACTTCCATGCAAG 420
 DB 361 CTTCCTCTTGGGAGTGTGATGCGTGGCAAACTGTACTTTGGGAACTTCCATGCAAG 420
 QY 421 GCAGTCCATGTCATCTACACAGTCAACCTCTACAGCACTGTCCTATCTGGCTTATC 480
 421 GCAGTCCATGTCATCTACACAGTCAACCTCTACAGCACTGTCCTATCTGGCTTATC 480
 DB 421 GCAGTCCATGTCATCTACACAGTCAACCTCTACAGCACTGTCCTATCTGGCTTATC 480
 QY 481 AGTGTGACCCGCTACCTGGCCATGTCACGCGCACCAACATCAGAGGCCAAGAACTG 540
 481 AGTGTGACCCGCTACCTGGCCATGTCACGCGCACCAACATCAGAGGCCAAGAACTG 540
 DB 481 AGTGTGACCCGCTACCTGGCCATGTCACGCGCACCAACATCAGAGGCCAAGAACTG 540
 QY 541 TTGCTGTAAGAGTGTCTATGTTGGCTGTGATCCCTGCCCTCTCTGACTATTCCTC 600
 541 TTGCTGTAAGAGTGTCTATGTTGGCTGTGATCCCTGCCCTCTCTGACTATTCCTC 600
 DB 541 TTGCTGTAAGAGTGTCTATGTTGGCTGTGATCCCTGCCCTCTCTGACTATTCCTC 600
 QY 601 GACTTCATCTTTGCCAACGTCAGTACGAGTACAGATATATCTGTGACCGCTTCTAC 660
 601 GACTTCATCTTTGCCAACGTCAGTACGAGTACAGATATATCTGTGACCGCTTCTAC 660
 DB 601 GACTTCATCTTTGCCAACGTCAGTACGAGTACAGATATATCTGTGACCGCTTCTAC 660
 QY 661 CCCAATGCTGTGGGAGTGTGTTCCAGTTTCAGACATCATGTTGGCTTATCCCTG 720
 661 CCCAATGCTGTGGGAGTGTGTTCCAGTTTCAGACATCATGTTGGCTTATCCCTG 720
 DB 661 CCCAATGCTGTGGGAGTGTGTTCCAGTTTCAGACATCATGTTGGCTTATCCCTG 720
 QY 721 CCTGATATGTCATCTCTCTCTCTATGTCATATCATCTCCAAAGCTGTACATCTCCAG 780
 721 CCTGATATGTCATCTCTCTCTCTATGTCATATCATCTCCAAAGCTGTACATCTCCAG 780
 DB 721 CCTGATATGTCATCTCTCTCTCTATGTCATATCATCTCCAAAGCTGTACATCTCCAG 780
 QY 781 GGGCACCAGAAAGCGCAAGGCCCTCAAGACACAGTATCTCTATCTCTGCTTTCTGCGC 840
 781 GGGCACCAGAAAGCGCAAGGCCCTCAAGACACAGTATCTCTATCTCTGCTTTCTGCGC 840
 DB 781 GGGCACCAGAAAGCGCAAGGCCCTCAAGACACAGTATCTCTATCTCTGCTTTCTGCGC 840
 QY 841 TGTGGCTGCTTACTATCATCTGGATCGACATCGACTCTTATCTCTCTGGAATATC 900
 841 TGTGGCTGCTTACTATCATCTGGATCGACATCGACTCTTATCTCTCTGGAATATC 900
 DB 841 TGTGGCTGCTTACTATCATCTGGATCGACATCGACTCTTATCTCTCTGGAATATC 900
 QY 901 AAGCAAGGATGTAGTGTGAGAACACTGTGCAAGATGATTCATCCAGAGGCCCTTA 960
 901 AAGCAAGGATGTAGTGTGAGAACACTGTGCAAGATGATTCATCCAGAGGCCCTTA 960
 DB 901 AAGCAAGGATGTAGTGTGAGAACACTGTGCAAGATGATTCATCCAGAGGCCCTTA 960
 QY 961 GCTTCTTCTCACTGTGTGTGAACCCATCTCTATGCTTTCTTGGAGCCAAATTTTAA 1020
 961 GCTTCTTCTCACTGTGTGTGAACCCATCTCTATGCTTTCTTGGAGCCAAATTTTAA 1020
 DB 961 GCTTCTTCTCACTGTGTGTGAACCCATCTCTATGCTTTCTTGGAGCCAAATTTTAA 1020

QY 1021 ACCTTGCCAGACGACGACTCACTCTGTGACAGAGGGTCCAGCTCAAGATCTCTCC 1080
 1021 ACCTTGCCAGACGACGACTCACTCTGTGACAGAGGGTCCAGCTCAAGATCTCTCC 1080
 DB 1021 ACCTTGCCAGACGACGACTCACTCTGTGACAGAGGGTCCAGCTCAAGATCTCTCC 1080
 QY 1081 AAGGAAAGCGAGGTGGACATTCATCTGTTTCCATCTGACTCTGAGCTTCAAGTTTCC 1140
 1081 AAGGAAAGCGAGGTGGACATTCATCTGTTTCCATCTGACTCTGAGCTTCAAGTTTCC 1140
 DB 1081 AAGGAAAGCGAGGTGGACATTCATCTGTTTCCATCTGACTCTGAGCTTCAAGTTTCC 1140
 QY 1141 TCCAGTAAACAGATGTAAAGACTTTTATACGATTAATTAATCTTTTAAAGTT 1200
 1141 TCCAGTAAACAGATGTAAAGACTTTTATACGATTAATTAATCTTTTAAAGTT 1200
 DB 1141 TCCAGTAAACAGATGTAAAGACTTTTATACGATTAATTAATCTTTTAAAGTT 1200
 QY 1201 ACACATTTTTCAGATATAAAGACTACCAATATTTGATGTTTATTTGCTTGGAT 1260
 1201 ACACATTTTTCAGATATAAAGACTACCAATATTTGATGTTTATTTTATTTGAT 1260
 DB 1201 ACACATTTTTCAGATATAAAGACTACCAATATTTGATGTTTATTTTATTTGAT 1260
 QY 1261 TTTTGTCTGTGTTCTTTAGTTTGTGAGTTTAAATGACTTATTTATATAATTTT 1320
 1261 TTTTGTCTGTGTTCTTTAGTTTGTGAGTTTAAATGACTTATTTATATAATTTT 1320
 DB 1261 TTTTGTCTGTGTTCTTTAGTTTGTGAGTTTAAATGACTTATTTATATAATTTT 1320
 QY 1321 TTTGTTTCATATGATGATGTGTCTAGCAGACCTGTGSCCAAGTTCTTATGCTGTAT 1380
 1321 TTTGTTTCATATGATGATGTGTCTAGCAGACCTGTGSCCAAGTTCTTATGCTGTAT 1380
 DB 1321 TTTGTTTCATATGATGATGTGTCTAGCAGACCTGTGSCCAAGTTCTTATGCTGTAT 1380
 QY 1381 GTCTGCTGTAGGACGTGAGAAAGGGAAGTCAATTCAGAGCGGTGTGATTCAGC 1440
 1381 GTCTGCTGTAGGACGTGAGAAAGGGAAGTCAATTCAGAGCGGTGTGATTCAGC 1440
 DB 1381 GTCTGCTGTAGGACGTGAGAAAGGGAAGTCAATTCAGAGCGGTGTGATTCAGC 1440
 QY 1441 TAAAGCTAGAAAGATCCCGAGCTTTATGCAATAGTAATCTCTCATTCCTGGTGAAC 1500
 1441 TAAAGCTAGAAAGATCCCGAGCTTTATGCAATAGTAATCTCTCATTCCTGGTGAAC 1500
 DB 1441 TAAAGCTAGAAAGATCCCGAGCTTTATGCAATAGTAATCTCTCATTCCTGGTGAAC 1500
 QY 1501 GTTTTCTCTGTTCTTAAAGCTGATTTTGTGTGAGAAAGTGGCACTTATACCAAGCCC 1560
 1501 GTTTTCTCTGTTCTTAAAGCTGATTTTGTGTGAGAAAGTGGCACTTATACCAAGCCC 1560
 DB 1501 GTTTTCTCTGTTCTTAAAGCTGATTTTGTGTGAGAAAGTGGCACTTATACCAAGCCC 1560
 QY 1561 AAAGTGTATGAAATCTGTTTTTTCAGTTTCAGAGTGGGTGATTTTCAGACCTTAC 1620
 1561 AAAGTGTATGAAATCTGTTTTTTCAGTTTCAGAGTGGGTGATTTTCAGACCTTAC 1620
 DB 1561 AAAGTGTATGAAATCTGTTTTTTCAGTTTCAGAGTGGGTGATTTTCAGACCTTAC 1620
 QY 1621 AGTGTACAGTGTGTATTAAGTTGTATATAAAGTACATGTTAACTTAAATTAATTA 1680
 1621 AGTGTACAGTGTGTATTAAGTTGTATATAAAGTACATGTTAACTTAAATTAATTA 1680
 DB 1621 AGTGTACAGTGTGTATTAAGTTGTATATAAAGTACATGTTAACTTAAATTAATTA 1680
 QY 1681 AAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCCGACACACGCTGGAATTC 1737
 1681 AAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCCGACACACGCTGGAATTC 1737
 DB 1681 AAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCCGACACACGCTGGAATTC 1737

RESULT 4
 X15882
 ID X15882 standard; cDNA: 1944 BP.
 AC X15882:
 DT 12-MAY-1999 (first entry)
 DE cDNA encoding G-protein coupled receptor polypeptide designated CXCR4B.
 KW G-protein coupled receptor; CXCR4B; human; splice variant;
 KW chemokine receptor; CXCR4; vaccine; gene therapy; HIV-1; HIV-1;
 KW cancer; pain; diabetes; obesity; anorexia; bulimia; asthma;
 KW Parkinson's disease; acute heart failure; hypotension; hypertension;
 KW urinary retention; osteoporosis; angina pectoris; myocardial infarction;
 KW stroke; ulcer; allergy; benign prostatic hyperplasia; migraine;
 KW vomiting; psychotic disorder; neurological disorder; anxiety;
 KW schizophrenia; manic depression; delirium; dementia; mental retardation;
 KW dyskinesias; Huntingtons disease; Gilles de la Tourette syndrome; ss.
 OS Homo sapiens.
 FH Key location/Qualifiers
 FT CDS 336..1406
 FT misc_feature 361
 FT /*tag= a
 FT /*tag= b
 FT /*note= "splice acceptor site"
 PN EP-897980-A2.


```

RESULT 5
ID V18357 standard: DNA, 1317 BP.
AC V18357:
DT 25-SEP-1998 (first entry)
DE Human RM3 seven transmembrane (7TM) receptor cDNA.
KW V28; Placenta; seven transmembrane receptor; 7TM; signal transduction;
OS Homo sapiens; inflammation; RM3; ss.
FH key Location/Qualifiers
FT CDS 201..1211
FT FT /*tag= a
FT FT /product= "Human RM3 seven transmembrane receptor"
PD 02-JUN-1998.
PD 17-NOV-1993: 153848.
PR 17-NOV-1992: US-977452.
PI (ICOS-) ICOS CORP.
PI Godiska R, Gray PW, Schweickart VL;
PI WPI; 98-332132/29.
DR P-PSDB: W48734.
PT DNA encoding V28 seven transmembrane receptor polypeptide - useful
PT for producing recombinant polypeptide and anti-V28 antibodies, and
PT in screening assays for V28 agonists and antagonists
PS Example 11: Columns 89-94; 56bp; English.
CC The present novel sequence represents the human RM3 cDNA isolated
CC from a human macrophage cDNA library. The invention claims for a
CC the full length V28 genomic DNA sequence (V18343) isolated from a human
CC placenta genomic library. The V28 (W48722) and RM3 proteins are seven
CC transmembrane (7TM) receptors which are probably involved in signal
CC transduction. The invention also claims that cells transformed with
CC V28 DNA can be used to produce the recombinant polypeptide, to produce
CC anti-V28 antibodies or in screening assays for V28 agonists or
CC antagonists. The antibodies, agonists and antagonists could then be
CC used to modulate V28 receptor-ligand binding, for e.g. in immunological
CC and/or inflammatory events in vivo.
SQ Sequence 1317 BP: 352 A; 342 C; 265 G; 378 T;

```

```

Query Match 61.4%; Score 1067; DB 1; Length 1317;
Best local similarity 100.0%; Pred. No. 0;
Matches 1067; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 106 ATATACACTTCAGATACTACACCGAGAAATGGCTCAGGGGACTATGATCCATGAAG 165
DB 171 ATATACACTTCAGATACTACACCGAGAAATGGCTCAGGGGACTATGATCCATGAAG 230
QY 166 GAACCCGTTCCGTGAAGAAATGCTAATTGAATTAATCTTCCGCCACCATCTAC 225
DB 231 GAACCCGTTCCGTGAAGAAATGCTAATTGAATTAATCTTCCGCCACCATCTAC 290
QY 226 TCCATCATCTTCTTAAGTGGCATTTGGGCAATGATGGTTCATCCGCGCATGGGTTAC 285
DB 291 TCCATCATCTTCTTAAGTGGCATTTGGGCAATGATGGTTCATCCGCGCATGGGTTAC 350
QY 286 CAGAGAACTGAGAGCATGACGGACAGTACAGGCTGCACCTGTGAGTGGCCGACCTC 345
DB 351 CAGAGAACTGAGAGCATGACGGACAGTACAGGCTGCACCTGTGAGTGGCCGACCTC 410
QY 346 CTCTTTGTCATCAGCGCTCCCTTGGGCGAGTGGAGCCGTGGCAAACTGGTACTTTGGG 405
DB 411 CTCTTTGTCATCAGCGCTCCCTTGGGCGAGTGGAGCCGTGGCAAACTGGTACTTTGGG 470
QY 406 AACTCTCTATGCAAGGAGTGCATGTCATCTACACAGTCAACCTCTACAGCATGTCTTC 465
DB 471 AACTCTCTATGCAAGGAGTGCATGTCATCTACACAGTCAACCTCTACAGCATGTCTTC 530
QY 466 ATCTTGAGCTTTCATCAGTGTGACCGCTACCTTGGCCATGCTCCAGCCACCAAGTCAAG 525
DB 531 ATCTTGAGCTTTCATCAGTGTGACCGCTACCTTGGCCATGCTCCAGCCACCAAGTCAAG 590
QY 526 AGGCCAAGGAGTGTGGTGAAGGTGTGTATGTTGGCTCTGATCCCTGCCCTC 585

```

```

DB 591 AGGCCAAGGAGTGTGGTGAAGGTGTGTATGTTGGGCTGTGATCCCTGCCCTC 650
QY 586 CTGCTGACTATTTCCGACATTCATCTTTGCCAACGCTAGTGGGAGATGACAGATATATC 645
DB 651 CTGCTGACTATTTCCGACATTCATCTTTGCCAACGCTAGTGGGAGATGACAGATATATC 710
QY 646 TGTGACCGCTTACCCCAATGACTGTGGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 705
DB 711 TGTGACCGCTTACCCCAATGACTGTGGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 770
QY 706 GTTGCCCTTATTCCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 765
DB 771 GTTGCCCTTATTCCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 830
QY 766 CTGTACACTCCAGAGGCCACAGAGGCCAGGCCCTTAAGACACATATCTCTATC 825
DB 831 CTGTACACTCCAGAGGCCACAGAGGCCAGGCCCTTAAGACACATATCTCTATC 890
QY 826 CTGCTTCTTCCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 885
DB 891 CTGCTTCTTCCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 950
QY 886 CTCTGGAATATCATCAAGAGGCTGTGATTGAGACACTGTGCACAAAGTGAATTTCC 945
DB 951 CTCTGGAATATCATCAAGAGGCTGTGATTGAGACACTGTGCACAAAGTGAATTTCC 1010
QY 946 ATACCGAGGCTTACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1005
DB 1011 ATACCGAGGCTTACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1070
QY 1006 GAGGCCAATTTAAACCTCTGCCCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1065
DB 1071 GAGGCCAATTTAAACCTCTGCCCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1130
QY 1066 CTGAATATCTCTCCCAAGAGGAGGTGACATTCATCTGTTTCCACTGAGTCTGAG 1125
DB 1131 CTGAATATCTCTCCCAAGAGGAGGTGACATTCATCTGTTTCCACTGAGTCTGAG 1190
QY 1126 TCTTCAAGTTTCACTCCAGCTACACAGATGTAAGACTTTT 1172
DB 1191 TCTTCAAGTTTCACTCCAGCTACACAGATGTAAGACTTTT 1237

```

```

RESULT 6
ID 066179 standard: cDNA, 1317 BP.
AC 066179:
DT 07-FEB-1995 (first entry)
DE Seven transmembrane receptor (RM3) coding sequence.
KW Primer; seven transmembrane receptor; receptor; amplification; PCR;
KW polymerase chain reaction; ss.
OS Homo sapiens.
FH key Location/Qualifiers
FT CDS 201..1214
FT FT /*tag= a
FT FT /product= "Seven transmembrane receptor."
PD WO9412635-A.
PD 09-JUN-1994.
PD 17-NOV-1993: U11153.
PR 17-NOV-1992: US-977452.
PI (ICOS-) ICOS CORP.
PI Godiska R, Gray PW, Schweickart VL;
PI WPI; 94-200264/24.
DR P-PSDB: R53753.
PT DNA encoding seven transmembrane receptors - used to develop
PT prods. for use as therapeutic or diagnostic agents for conditions
PT involving the receptors.
PS Example 11: Page 82-83; 100bp; English.
CC Two primers (066148, 066149) were used in a PCR reaction containing
CC human macrophage cDNA library in plasmid pRC/CMV. When the PCR
CC products were subjected to agarose gel electrophoresis a faint band
CC of 180-200 base pairs was observed. Re-amplified material was
CC digested with BamHI and HindIII and cloned into the plasmid

```

CC Bluescript SK-. Of sixteen clones sequenced, two contained a unique
CC sequence termed RM3. Specific primers for the partial RM3 clone were
CC used to identify this full length RM3 CDNA clone.
SQ Sequence 1317 BP: 332 A: 343 C: 264 G: 378 T:

Query Match 58.5%; Score 1016; DB 1; Length 1317;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1056; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 105 ATATACACTTCAGATAACTACACCGAGGAATGGGCTCAGGGAGTACTGATCCATGAG 165
DB 171 ATATACACTTCAGATAACTACACCGAGGAATGGGCTCAGGGAGTACTGATCCATGAG 230
QY 166 GAACCCGTGTTCCGTGAGAAATGCTAATTTCAATAAATCTTCTGCCACCATTAC 225
DB 231 GAACCCGTGTTCCGTGAGAAATGCTAATTTCAATAAATCTTCTGCCACCATTAC 290
QY 226 TCATCATCTTCTTAAGTGGCAATGGGCAATGGATGGATGCTGATGGGTAC 285
DB 291 TCATCATCTTCTTAAGTGGCAATGGGCAATGGATGGATGCTGATGGGTAC 350
QY 286 CAGAGAACTGAGAAAGATGAGGAGACAGTACAGGCTGACCTGTGATGGCGACCTC 345
DB 351 CAGAGAACTGAGAAAGATGAGGAGACAGTACAGGCTGACCTGTGATGGCGACCTC 410
QY 346 CTCTTTGTATCATACGCTTCCCTTCTGGGCAAGTGTGATCCGTGCAAACTGGTATGGG 405
DB 411 CTCTTTGTATCATACGCTTCCCTTCTGGGCAAGTGTGATCCGTGCAAACTGGTATGGG 470
QY 406 AACTTCATGACAGGAGGATGCTATGATACAGTCAAGCTCTACAGAGTGTCTC 465
DB 471 AACTTCATGACAGGAGGATGCTATGATACAGTCAAGCTCTACAGAGTGTCTC 530
QY 466 ATCTGGGCTTCAATGAGTGTGAGCCGTACCTGGCCATGCTCCAGCGCACCAACAGTACG 525
DB 531 ATCTGGGCTTCAATGAGTGTGAGCCGTACCTGGCCATGCTCCAGCGCACCAACAGTACG 590
QY 526 AGGCAAGGAGGATGTTGGTGAAGAGTGTGATGTTGGGCTGTGATGCTGCTCC 585
DB 591 AGGCAAGGAGGATGTTGGTGAAGAGTGTGATGTTGGGCTGTGATGCTGCTCC 650
QY 586 CTGCTGACTATTCGCGACTTCACTTCTTGCCACAGTCAAGTGAAGGAGTCAATATATC 645
DB 651 CTGCTGACTATTCGCGACTTCACTTCTTGCCACAGTCAAGTGAAGGAGTCAATATATC 710
QY 646 TGTGACCGCTTACCCCAATGACTGTGGTGGTGTGTTCCAGTTTCAAGCATATC 705
DB 711 TGTGACCGCTTACCCCAATGACTGTGGTGGTGTGTTCCAGTTTCAAGCATATC 770
QY 706 GTGGGCTTATTCGCTGCTGATTTGTATGCTGTCGCTGATTTGATCATCTCCAG 765
DB 771 GTGGGCTTATTCGCTGCTGATTTGTATGCTGTCGCTGATTTGATCATCTCCAG 830
QY 766 CTGTCACACTTCAAGGAGGAGCCAGCAAGGAGGAGGCTTCAAGCATATCTTCATC 825
DB 831 CTGTCACACTTCAAGGAGGAGCCAGCAAGGAGGAGGCTTCAAGCATATCTTCATC 890
QY 826 CTGGCTTCTTCCGCTGTGGCTGTGCTTACTAATTTGGATTCAGCATCTCTTCATC 885
DB 891 CTGGCTTCTTCCGCTGTGGCTGTGCTTACTAATTTGGATTCAGCATCTCTTCATC 950
QY 886 CTCTGGAATCATCAAGCAAGGAGTGTGATTTGAGAACTGTGCAAGTGGATTTCC 945
DB 951 CTCTGGAATCATCAAGCAAGGAGTGTGATTTGAGAACTGTGCAAGTGGATTTCC 1010
QY 946 ATACCGAGGAGGCTAGCTTCTTCACTGTGTTGTGAAGCCCATCTCTATGCTTTCTT 1005
DB 1011 ATACCGAGGAGGCTAGCTTCTTCACTGTGTTGTGAAGCCCATCTCTATGCTTTCTT 1070
QY 1006 GAGGCCAAATTTAAAGCTCTGCGCAGACGACACTACCTCTGTGAGCAAGGCTCCAGC 1065
DB 1071 GAGGCCAAATTTAAAGCTCTGCGCAGACGACACTACCTCTGTGAGCAAGGCTCCAGC 1130

QY 1066 CTCAGATCTCTTCCAAAGGAGGAGTGTGACATTCATCTGTTCCAGTGTGAG 1125
DB 1131 CTCAGATCTCTTCCAAAGGAGGAGTGTGACATTCATCTGTTCCAGTGTGAG 1190
QY 1126 TCTTCAAGTTTTCACCTCCACTACACAGATGTAAGACTTTT 1172
DB 1191 TCTTCAAGTTTTCACCTCCACTACACAGATGTAAGACTTTT 1237

RESULT 7

ID X15883 standard; cDNA: 611 BP.
AC X15883;
DT 12-MAY-1999 (first entry)
DE cDNA encoding a partial CXCR4B protein.
KW G-protein coupled receptor; CXCR4B; human; splice variant;
KW chemokine receptor; CXCR4; vaccine; gene therapy; HIV-1;
KW cancer; pain; diabetes; obesity; anorexia; bulimia; asthma;
KW Parkinson's disease; acute heart failure; hypotension; hypertension;
KW urinary retention; osteoporosis; angina pectoris; myocardial infarction;
KW stroke; ulcer; allergy; benign prostatic hypertrophy; migraine;
KW vomiting; psychotic disorder; neurological disorder; anxiety;
KW schizophrenia; manic depression; delirium; dementia; mental retardation;
KW dyskinesias; Huntingtons disease; Gilles de la Tourette syndrome; ss.
OS Homo sapiens.
PN EP-897980-A2.
PD 24-FEB-1999.
PF 07-AUG-1998; 306324.
PR 24-JUL-1998; US-056601.
PR 20-AUG-1997; US-056601.
PA (SMK) SMITHKLINE BEECHAM CORP.
PI Gupta SK, Pillariseti K.
DR WPI: 99-134643/12.
DR P-PDB: W97363.
PT New G protein coupled receptor (CXCR4B) polypeptide and
PT polynucleotide, human splice variant of a chemokine receptor -
PT useful as diagnostic reagents and for prevention and treatment of
PT HIV infection, cancer, stroke and dementia
PS Claim 13; Page 18; 24pp; English.
CC The present sequence encodes a partial G-protein coupled receptor
CC designated CXCR4B, which is a human splice variant of a chemokine
CC receptor CXCR4. CXCR4B polynucleotides and polypeptides are useful for
CC diagnosing susceptibility to diseases by detecting mutations or
CC polymorphisms in the CXCR4B gene or analysing for the presence or amount
CC of CXCR4B polypeptide expressed in a patient. CXCR4B polypeptides and
CC polynucleotides are also useful for screening for antagonists and
CC agonists which can be used to treat conditions associated with CXCR4B
CC polypeptide imbalance. CXCR4B polypeptides can be administered directly
CC (as a vaccine) or via a vector (gene therapy) to prevent disease.
CC Diseases diagnosed, prevented and treated include: HIV-1 and HIV-1
CC infections; cancer; pain; diabetes; obesity; anorexia; bulimia; asthma;
CC Parkinson's disease; acute heart failure; hypotension; hypertension;
CC urinary retention; osteoporosis; angina pectoris; myocardial infarction;
CC stroke; ulcers; allergies; benign prostatic hypertrophy; migraine;
CC vomiting; psychotic and neurological disorders, including anxiety,
CC schizophrenia, manic depression, delirium, dementia and severe mental
CC retardation; and dyskinesias, such as Huntingtons disease or Gilles de
CC la Tourette syndrome. CXCR4B polypeptides are useful for mapping genes
CC to chromosomes, allowing gene inheritance to be studied through linkage
CC analysis.
SQ Sequence 611 BP: 166 A: 125 C: 139 G: 181 T:

Query Match 14.3%; Score 249; DB 1; Length 611;
Best Local Similarity 100.0%; Pred. No. 3e-80;
Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 ATATACACTTCAGATAACTACACCGAGGAATGGGCTCAGGGAGTACTGATCCATGAG 165
DB 363 ATATACACTTCAGATAACTACACCGAGGAATGGGCTCAGGGAGTACTGATCCATGAG 422
QY 166 GAACCCGTGTTCCGTGAGAAATGCTAATTTCAATAAATCTTCTGCCACCATTAC 225


```

Db 423 GAACCTGTTCCGAGAAAAATGTAATTTCAATTAATTTCTCTGCCACACATCTAC 482
Oy 226 TCCATCATCTTTCTTAAGTGGCAATGATTTGGTATGCTGTCATGGGTTAC 285
Db 483 TCCATCATCTTTCTTAAGTGGCAATGATTTGGGCAATGATTTGGTATGCTGTCATGGGTTAC 542
Oy 286 CAGAGAACTAGAGAGCATGACGACAGTACAGGCTGACCTGACATGGCCGACCTC 345
Db 543 CAGAGAACTAGAGAGCATGACGACAGTACAGGCTGACCTGACATGGCCGACCTC 602
Oy 346 CTCTTTGTC 354
Db 603 CTCTTTGTC 611

RESULT 8
T20146
ID T20146 standard; cDNA to mRNA; 218 BP.
AC T20146.
DT 31-JUL-1996 (first entry)
DE Human gene signature HUMGS01292.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
human; cloning; mapping; non-biased library; diagnosis; detection;
cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN W09514772.A1.
PD 01-JUN-1995.
PE 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
WP1: 95-206931/27.
DR Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1: Page 571-572; 2245pp: Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in 119001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues: synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 218 BP; 61 A; 38 C; 44 G; 75 T;

Query Match 12.6%; Score 218; DB 1; Length 218;
Best Local Similarity 100.0%; Pred. No. 3.9e-69; Indels 0; Gaps 0;
Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy 1634 GTATTAGTTGTAATAAAAGTACATGTTAACTTAA 1671
Db 181 GTATTAGTTGTAATAAAAGTACATGTTAACTTAA 218

RESULT 9
O44391
ID O44391 standard; cDNA to mRNA; 3581 BP.
AC O44391.
DT 14-SEP-1994 (first entry)
DE Sequence of murine OSF-4 cDNA.
KW OSF-4; cadherin; growth factor; osteogenesis; osteoblast; therapy;
KW diagnosis; ss.
OS Mus musculus.
FH Key location/Qualifiers
FT cds 284..2671
FT /*tag= a

PN EP-585801-A.
PD 09-MAR-1994.
PE 25-AUG-1993; J13602.
PR 28-AUG-1992; JP-230028.
PA (FARH) HOECHST JAPAN LTD.
PI Amann E, Kawai S, Okazaki M, Takeshita S, Tsujimura A;
DR WP1: 94-076132/10.
DR P-PSDB: R49730.
PT New bone related, cadherin-like OSF-4 proteins - for treatment
PT and diagnosis of bone metabolic disease, and nucleic acid
PT encoding them
PS Claim 3: Page 13-17; 34pp: English.
CC cDNA libraries were constructed from the mouse osteoblastic cell
CC line MC3T3-E1 and from mouse liver tissue, amplified by PCR, and
CC then as much common DNA as possible removed by hybridisation between
CC the 2 libraries. Residual E1-specific DNA was amplified, inserted
CC into lambda gt10 and screened by plaque hybridisation. A minibank of
CC 273 E-specific clones was recovered, their inserts amplified and
CC used to screen total RNA from both cell types. One clone specific
CC for E1 was identified and sequenced. The insert from this clone was
CC used to screen cDNA prep. from E1 RNA and the longest posn. insert
CC cloned in pGEM 112f (+) to give pKOT164. This insert was sequenced;
CC it encoded the 796 AA mouse precursor protein (O44391/R49730). The
CC insert was also used to screen a cDNA bank prep. from human
CC osteosarcoma to identify 2 clones encoding the 2 human precursor
CC proteins - OSF-4-1 and OSF-4-2 (O44392/R49731 and O44393/R49732
CC respectively).
SQ Sequence 3581 BP; 1121 A; 774 C; 800 G; 886 T;

Query Match 3.1%; Score 53; DB 1; Length 3581;
Best Local Similarity 100.0%; Pred. No. 9.3e-11; Indels 0; Gaps 0;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1669 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGCGCCG 1721
Db 3523 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGCGCCG 3575

RESULT 10
O65607
ID O65607 standard; cDNA; 1338 BP.
AC O65607.
DT 31-JUN-1995 (first entry)
DE Rabbit zona pellucida ZPC coding sequence.
KW Rabbit; lapine; zona pellucida; ZPC; immunocontraception; ds.
OS Oryctolagus cuniculus.
FH key location/Qualifiers
FT cds 17..1264
FT /*tag= a
FT /product= ZPC

PN W09411019-A.
PD 26-MAY-1994.
PE 06-NOV-1993; U10851.
PR 09-NOV-1992; US-973341.
PR 29-JUN-1993; US-012990.

```



```

PA (ZONA-) ZONAGEN INC.
PI Harris JD, Hsu KT, Podolski JS;
DR WPI: 94-183156/22.
DR P-PSDB: R55197.
PI Use of zona pellucida proteins and antibodies - for inducing
PT reproducible transient infertility or permanent sterility in
PT female mammals
PS Claim 22: Page 81-83; 154pp; English.
CC A cDNA library was prepared from mRNA isolated from ovaries removed
CC from 5 week old rabbits. The lambda gt10 library was screened with
CC a porcine ZPC cDNA probe (Q055606). Two positive clones were sequenced
CC and were found to be identical except that one contained 4 extra
CC nucleotides at the 5'-end. The determined sequence was 75%
CC homologous to DNA coding for porcine ZPC.
SO Sequence 1338 BP; 255 A; 445 C; 402 G; 236 T;

Query Match 3.1%; Score 53; DB 1; Length 1338;
Best Local Similarity 100.0%; Pred. No. 1.1e-10;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1559 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCGC 1721
DB 1280 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCGC 1332

RESULT 11
V05740
ID V05740 standard; cDNA: 87 BP.
AC V05740.
DT 05-JUN-1998 (first entry)
DE Nucleotide sequence of the 3' portion of clone AS301_2.
KW Secreted protein; homology; antibody; immunoassay reagent;
KW nutritional supplement; therapeutic activity; ds.
OS Homo sapiens.
PN WO9746683-A2.
PD 11-DEC-1997.
PF 06-JUN-1997; 009878.
PR 07-JUN-1996; US-659224.
PA (GEMV ) GENETICS INST INC.
PI Bowman M, Evans C, Jacobs K, Lavallic ER, McCly JM,
PI Merberg D, Racie LA, Spaulding V, Treacy M;
DR WPI: 98-042191/04.
PT Nucleic acids encoding secreted proteins from clones within ATCC
PT 98076 - useful as immuno-modulators, anti-proliferative agents,
PT regulators of cell differentiation and tissue growth, etc
PS Claim 40: Page 76; 99pp; English.
CC The present sequence represents the nucleotide sequence of the 3'
CC portion of clone AS301_2. The clone was isolated from a human fetal
CC brain cDNA library using probe V05760. AS301_2 is a full length clone
CC encoding a secreted protein. V05739 provides 5' portion of the nucleotide
CC sequence, and V05738 an internal sequence. The nucleic acid can be used
CC for expression of recombinant proteins, as tissue, molecular weight or
CC chromosome markers, indicators of genetic disorders and sources
CC of probes and primers. They can also be used to generate anti-protein
CC or anti-DNA antibodies and as components of interaction trap assays etc.
CC The protein is useful for raising antibodies, as immunoassay reagents
CC and as nutritional supplements. The protein may possibly have any of a
CC great variety of therapeutic activities.
SO Sequence 87 BP; 79 A; 4 C; 4 G; 0 U;

Query Match 3.1%; Score 53; DB 1; Length 87;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1559 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCGC 1721
DB 35 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCGC 87

RESULT 12
V05728

```

```

ID V05728 standard; cDNA: 94 BP.
AC V05728.
DT 05-JUN-1998 (first entry)
DE Nucleotide sequence of the 3' portion from clone A1147_1.
KW Secreted protein; antibody; immunoassay reagent;
KW nutritional supplement; therapeutic activity; murine; calmegin; ds.
OS Homo sapiens.
PN WO9746683-A2.
PD 11-DEC-1997.
PF 06-JUN-1997; 009878.
PR 07-JUN-1996; US-659224.
PA (GEMV ) GENETICS INST INC.
PI Bowman M, Evans C, Jacobs K, Lavallic ER, McCly JM,
PI Merberg D, Racie LA, Spaulding V, Treacy M;
DR WPI: 98-042191/04.
PT Nucleic acids encoding secreted proteins from clones within ATCC
PT 98076 - useful as immuno-modulators, anti-proliferative agents,
PT regulators of cell differentiation and tissue growth, etc
PS Claim 25: Page 67; 99pp; English.
CC The present sequence represents the nucleotide sequence of the 3'
CC portion of clone A1147_1. The clone was isolated from a human adult
CC testes cDNA library using probe V05755. A1147_1 is a full length
CC clone encoding a secreted protein. V05727 provides the 5' portion
CC sequence. A1147_1 shows some homology with murine calmegin, a
CC Ca2+-binding protein that is specifically expressed in spermatogenesis.
CC A1147_1 may share some of its activity. The nucleic acid can be used for
CC expression of recombinant proteins, as tissue, molecular weight or
CC chromosome markers, indicators of genetic disorders and sources of probes
CC and primers. They can also be used to generate anti-protein or anti-DNA
CC antibodies and as components of interaction trap assays etc. The protein
CC is useful for raising antibodies, as immunoassay reagents and as
CC nutritional supplements. The protein may possibly have any of a great
CC variety of therapeutic activities.
SO Sequence 94 BP; 86 A; 4 C; 4 G; 0 U;

Query Match 3.1%; Score 53; DB 1; Length 94;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1569 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCGC 1721
DB 42 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCGC 94

RESULT 13
V61487
ID V61487 standard; cDNA: 4237 BP.
AC V61487.
DT 11-JAN-1999 (first entry)
DE Human secreted protein fe366.1 cDNA.
KW Secreted protein; human; fe366_1; ds.
OS Homo sapiens.
PN key location/Qualifiers
FT CDS 3746..4030
FT /tag= a
FT WO9841539-A2.
PD 24-SEP-1998.
PF 19-MAR-1998; 005474.
PR 18-MAR-1998; US-040963.
PR 19-MAR-1997; US-820493.
PA (GEMV ) GENETICS INST INC.
PI Agostino MJ, Jacobs K, Lavallic ER, McCoy JM, Merberg D,
PI Racie LA, Spaulding V, Treacy M;
DR WPI: 98-521163/44.
DR P-PSDB: W79096.
PI New polynucleotide(s) encoding secreted human proteins - derived
PI from human foetal kidney, adult testes and adult or foetal brain
PI cDNA libraries
PS Claim 36: Page 88-91; 112pp; English.
CC This full-length cDNA clone, designated fe366.1, codes for a novel
CC secreted human protein (see W79096). It was isolated from a human
CC adult brain cDNA library using methods which are selective for

```

```
CC cDNAs encoding secreted proteins, or was identified as encoding
CC a secreted or transmembrane protein on the basis of computer
CC analysis of the encoding protein. The nucleotide sequence shows
CC homology to some database sequences, and may contain a CAA repeat
CC and/or Alu repetitive element. The invention provides cDNA clones
CC (see V61477-87) from human foetal kidney, adult testis, and adult
CC foetal brain cDNA libraries that code for secreted proteins
CC (see W79087-97). These clones are deposited as ATCC 98364. The
CC polynucleotides and proteins are predicted to have useful
CC biological activities which would make them suitable for treating,
CC preventing or ameliorating medical conditions in humans and
CC animals, although no supporting data is given. Suggested
CC activities include nutritional, immune stimulating (e.g. as
CC vaccines) or suppressing, haematopoiesis regulating, tissue growth,
CC activin/inhibin, chemotactic/chemokinetic, haemostatic and
CC thrombolytic, receptor/ligand, antiinflammatory, cadherin/tumour
CC invasion suppressor and tumour inhibition activities. The
CC polynucleotides are also stated to be useful for gene therapy.
CC Sequence 4237 BP; 1330 A; 778 C; 784 G; 1338 T;
SQ

Query Match
Best Local Similarity 100.0%; Score 53; DB 1; Length 4237;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1669 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCCG 1721
DB 2554 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCCG 2606

RESULT 14
V64789
ID V64789 standard; CDNA; 1338 BP.
AC V64789;
DE 29-JAN-1999 (first entry)
KW Rabbit ZPC CDNA.
KW ZPC; zona pellucida; infertility; sterility; immunocontraceptive;
OS Oryctolagus cuniculus.
FH Key Location/Qualifiers
FT CDS 17..1264
FT /tag= a
FT /product= "zpc"
PN US5837497-A.
PD 17-NOV-1998.
PF 07-JUN-1995; 484993.
PR 09-NOV-1993; US-149223.
PR 09-NOV-1992; US-973441.
PR 29-JAN-1993; US-012990.
PR 07-JUN-1995; US-484993.
PA (ZONA-) ZONAGEN INC.
PI Harris JD;
PI WPI: 99-023447/02.
DR P-SPB: W81807.
PT Isolated zona pellucida DNA from different mammals - used to develop
PT products which can be used for vaccination to induce transient
PT infertility or permanent sterility in female mammals
PS Claim 1: Column 59-62; 84pp; English.
CC This sequence encodes a rabbit ZPC protein isolated from zona pellucida.
CC This protein can be used in a method for specifically inducing transient
CC infertility or permanent sterility in a host animal by selective
CC vaccination with specific zona pellucida proteins or
CC immunocontraceptively active fragments.
CC Sequence 1338 BP; 255 A; 445 C; 402 G; 236 T;
SQ
```

```
RESULT 15
V84799
ID V84799 standard; CDNA; 2744 BP.
AC V84799;
DE 30-MAR-1999 (first entry)
DE Nucleotide sequence encoding human prohormone convertase 4.
KW ds: human; prohormone convertase; testicular prohormone; testicular cell;
KW fertility; spermatogenesis; embryo development; chromosome 19; 19p13.3.
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 61..2328
FT /tag= a
FT /product= "prohormone convertase 4"
PN M09650560-A1.
PD 12-NOV-1998.
PF 01-MAY-1998; U08871.
PR 06-MAY-1997; US-044015.
PA (ZYMO ) ZYMOGENETICS INC.
PI Jaspers SR, Lok S;
PI WPI: 99-059673/05.
DR P-PSDB: W81365.
DE New nucleic acid encoding human prohormone convertase 4 - useful
DE for, e.g. identifying modulators and new testicular prohormones, and
DE diagnosing chromosomal abnormalities
PT Claim 2: Page 70-75; 95pp; English.
PS Human prohormone convertase 4 can be used to raise antibodies; for in
PS vitro identification of modulators and to identify, or determine
PS function of, (new) testicular prohormones, and for processing of such
PS prohormones (e.g. for stimulating proliferation or differentiation of
PS testicular cells). mRNA for PC4 is detected only in testis, suggesting
CC its involvement in fertility, spermatogenesis and early embryo
CC development. The nucleic acid and its fragments (particularly probes and
CC primers) are used to detect abnormalities in chromosome 19 (the gene for
CC PC4 is located at 19p13.3).
CC Sequence 2744 BP; 552 A; 961 C; 833 G; 398 T;
SQ
```

```
Query Match
Best Local Similarity 100.0%; Score 53; DB 1; Length 2744;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1669 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCCG 1721
DB 2692 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCCG 2744

RESULT 16
Q78635/C
ID Q78635 standard; DNA; 61 BP.
AC Q78635;
DE 14-JUL-1995 (first entry)
DE Human chondromodulin-1 cDNA primer.
KW Chondromodulin-1; chondrocyte growth regulation;
KW vascular endothelial cell growth; primer; ss.
OS Synthetic.
PN EP-624645-A.
PD 17-NOV-1994.
PF 11-MAY-1994; 107364.
PR 11-MAY-1993; JP-109620.
PR 17-DEC-1993; JP-318298.
PA (SUZU) SUZUKI F.
PA (MITU) MITSUBISHI KASEI CORP.
PI Hiraki Y, Kohara A, Kondo J, Mori A, Suzuki F;
PI Suzuki Y, Takahashi K, Yamada E;
DR WPI: 94-350785/44.
DE Human chondromodulin-1 protein - stimulates growth of
DE chondrocytes and inhibits vascular endothelial cell growth
PS Example 4: Page 10; 34pp; English.
CC Q78635 and Q78636 are a pair of primers for Q78627-Q78632 which
CC encode R65486-R65491 respectively, human chondromodulin-1 (CM-1),
CC a chondrocyte growth regulator. CM-1 as part of a pharmaceutical
```

CC compen. with a carrier, excipient or solvent, can be used to
CC stimulate chondrocyte growth. CM-1 also inhibits vascular
CC endothelial cell growth, by promoting the differential
CC potency of chondrocytes.
SO Sequence 61 BP: 2 A: 8 C: 8 G: 43 T:

Query Match 2.9%; Score 51; DB 1; Length 51;
Best Local Similarity 100.0%; Pred. No. 1,1e-09;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1673 AAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCCCA 1723
DB 61 AAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCCCA 11

RESULT 17
X26148
ID X26148 standard; CDNA: 775 BP.
AC X26148;
DT 21-MAY-1999 (first entry)
DE Human pan-s/tk-1A receptor polypeptide encoding CDNA.
KW pancreatic serine/threonine kinase receptor; pan-s/tk; cell growth;
KW differentiation; gene therapy; ligand; screening; cell proliferation;
KW regulator; pancreatic; ds.
OS Homo sapiens.
PN MO9907854-A2.
PD 18-FEB-1999.
PE 11-AUG-1998; U16640.
PF 11-AUG-1997; US-909354.
PR 11-AUG-1997; US-909354.
PT (ONTO-) ONTOGENY INC.
PI Barker DD, Miao N, Pang K;
DR WPI: 99-167430/14.
P-PSDB: Y01496.
PT New pancreatic serine/threonine kinase (pan-s/tk) receptor
PT polypeptides - useful for modulating cell proliferation,
PT differentiation and survival in animals
PS Claim 28: Page 66: 77pp; English.
CC The invention relates to isolated and/or recombinant pancreatic serine/
CC threonine kinase receptor (pan-s/tk) polypeptides. Sequences of rat and
CC human pan-s/tk polypeptides and nucleic acids encoding them are provided.
CC Host cells containing vectors comprising the pan-s/tk genes are
CC used for the recombinant expression of the proteins. The pan-s/tk
CC polypeptides are useful for modulating cell growth, differentiation or
CC survival in an animal cell. The polynucleotides are also useful for
CC detecting pan-s/tk ligand on cells in samples and for gene therapy.
CC Ligands identified, together with soluble polypeptides, are useful for
CC screening test compounds that modulate the bioactivity of a pan-s/tk
CC receptor. The discovery of this new class of pan-s/tk polypeptides
CC provides an alternative method of mediating cell proliferation,
CC differentiation and survival in animals. The polypeptides also represent
CC the first identified specific regulators of pancreatic development. The
CC present sequence represents a CDNA encoding a human pan-s/tk-1A receptor
CC polypeptide.
SO Sequence 775 BP: 231 A: 188 C: 179 G: 177 T:

Query Match 2.9%; Score 51; DB 1; Length 775;
Best Local Similarity 100.0%; Pred. No. 6,6e-10;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1653 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1713
DB 715 AACCTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 766

RESULT 18
X26151
ID X26151 standard; CDNA: 1718 BP.
AC X26151;
DT 21-MAY-1999 (first entry)
DE 3' CDNA sequence of human pan-s/tk-2 receptor polypeptide.
KW Pancreatic serine/threonine kinase receptor; pan-s/tk; cell growth;

KW differentiation; gene therapy; ligand; screening; cell proliferation;
KW regulator; pancreatic; ds.
OS Homo sapiens.
PN MO9907854-A2.
PD 18-FEB-1999.
PE 11-AUG-1998; U16640.
PF 11-AUG-1997; US-909354.
PR 11-AUG-1997; US-909354.
PT (ONTO-) ONTOGENY INC.
PI Barker DD, Miao N, Pang K;
DR WPI: 99-167430/14.

PT New pancreatic serine/threonine kinase (pan-s/tk) receptor
PT polypeptides - useful for modulating cell proliferation,
PT differentiation and survival in animals
PS Claim 28: Page 73: 77pp; English.
CC The invention relates to isolated and/or recombinant pancreatic serine/
CC threonine kinase receptor (pan-s/tk) polypeptides. Sequences of rat and
CC human pan-s/tk polypeptides and nucleic acids encoding them are provided.
CC Host cells containing vectors comprising the pan-s/tk genes are
CC used for the recombinant expression of the proteins. The pan-s/tk
CC polypeptides are useful for modulating cell growth, differentiation or
CC survival in an animal cell. The polynucleotides are also useful for
CC detecting pan-s/tk ligand on cells in samples and for gene therapy.
CC Ligands identified, together with soluble polypeptides, are useful for
CC screening test compounds that modulate the bioactivity of a pan-s/tk
CC receptor. The discovery of this new class of pan-s/tk polypeptides
CC provides an alternative method of mediating cell proliferation,
CC differentiation and survival in animals. The polypeptides also represent
CC the first identified specific regulators of pancreatic development. The
CC present sequence represents the 3' CDNA sequence of a human pan-s/tk-2
CC receptor polypeptide.
SO Sequence 1718 BP: 495 A: 335 C: 394 G: 492 T:

Query Match 2.9%; Score 51; DB 1; Length 1718;
Best Local Similarity 100.0%; Pred. No. 5,5e-10;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1653 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1713
DB 1659 AACCTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1709

RESULT 19
T74484
ID T74484 standard; CDNA: 495 BP.
AC T74484;
DT 01-NOV-1997 (first entry)
DE Pilosulin 2 precursor CDNA clone Myr p2.
KW Pilosulin 2; Australian jumper ant; venom; allergen; hypotensive;
KW blood pressure; cell growth; ss.
OS Myrmecia pilosula.
FH Key Location/Qualifiers
FT met_peptide 44..271
FT signal_peptide 44..121
FT mat_peptide 188..268
FT /*tag= a
FT /*tag= b
FT /*tag= c
FT /*product= pilosulin A
PN MO9713854-A1.
PD 17-APR-1997.
PE 09-OCT-1996; AU0632.
PR 09-OCT-1995; AU-005839.
PA (NSYD-) NORTHERN SYDNEY AREA HEALTH SERVICE.
PI Baldo B, Donovan G;
DR WPI: 97-235890/21.
DR P-PSDB: W21784.
PT New isolated Pilosulin 2 polypeptide - obtained from Myrmecia
PT Pilosula ant venom, used for reducing blood pressure in animals or
PT for inhibiting cell growth
PS Claim 1: Fig 1; 27pp; English.
CC This sequence is an isolated CDNA clone, designated Myr p2, which
CC includes a sequence encoding a polypeptide that has the biological

CC activity of pilosulin 2 having the amino acid sequence from
 CC residues 49-75 of a 75-amino acid sequence (W21784). It was
 CC isolated from a cDNA library prepared from the abdomens of Myrmecia
 CC pilosula ants by immunoscreening with sera from venom-allergic
 CC individuals. The isolated sequence can be incorporated into an
 CC expression vector and used to produce pilosulin 2 in transformed
 CC host cells. Pilosulin 2 and its analogues or fragments can be used
 CC in claimed methods for reducing the blood pressure of an animal,
 CC and for inhibiting the growth of (killing) cells.
 SO Sequence 495 BP; 201 A; 68 C; 99 G; 127 T;

Query Match 2.9%; Score 50; DB 1; Length 495;
 Best Local Similarity 100.0%; Pred. No. 1.6e-09;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1672 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGC 1721
 DB 446 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGC 495

RESULT 20

ID V00418 standard; cDNA: 201 BP.
 AC V00418: 1998 (first entry)
 DT 12-MAY-1998 (first entry)
 DE 3' fragment of clone M8_2.
 KW Human; secreted protein; molecular weight marker; genetic fingerprinting;
 KM antibody production; nutritional supplement; therapy; neural tissue;
 KW glioblastoma line T998G; clone M8_2; ds.
 OS Homo sapiens.
 PN M09740069.A2.
 PD 30-OCT-1997.
 PF 14-APR-1997: 006134.
 PR 19-APR-1996: US-635311.
 PA (GENY) GENETICS INST INC.
 PI Jacobs K, Lavallie ER, McCoy JM, Metberg D, Racie LA,
 PI Spaulding V:
 WP: 97-353776/49.
 DR Isolated nucleic acid clones from ATCC 98028 encode novel secreted
 PT proteins - having many potential uses, e.g. as immunomodulators,
 PT cell proliferation or differentiation inhibitors or hematopoiesis
 PT regulators
 PS Claim 22: Page 69: 114pp: English.
 CC This sequence represents the 3' end of clone M8_2, which is a
 CC polynucleotide of the invention. This sequence was isolated from a human
 CC neural tissue (glioblastoma line T998G) cDNA library. The polynucleotide,
 CC which encodes a secreted protein, can be used, e.g. as a tissue or
 CC molecular weight marker, in genetic fingerprinting, to raise anti-protein
 CC or anti-DNA antibodies and in interaction trap assays. The protein can be
 CC used to assay biological activity, raise antibodies for use in
 CC immunossays, as a marker, to identify inhibitors of its interactions and
 CC as a nutritional supplement. It may also have a very wide range of
 CC therapeutic and biological activities (no examples are given to support
 CC this), e.g. cytokine or modulator of cell proliferation and
 CC differentiation, immunostimulant or immunosuppressant, hematopoiesis
 CC regulator, bone, cartilage, tendon, ligament and/or nerve tissue growth
 CC stimulator, follicle inhibitor/stimulator, chemotactic/chemokinetic,
 CC haemostatic, thrombolytic or anti-inflammatory agent, antimicrobial,
 CC biohythm, metabolism or behaviour modifier, anti-depressant or analgesic
 CC or psoriasis treatative.
 SO Sequence 201 BP; 90 A; 16 C; 43 G; 39 T;

Query Match 2.9%; Score 50; DB 1; Length 201;
 Best Local Similarity 100.0%; Pred. No. 2e-09;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1672 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGC 1721
 DB 152 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGC 201

RESULT 21

ID V46370 standard; cDNA to mRNA; 1877 BP.
 AC V46370:
 DT 20-NOV-1998 (first entry)
 DE Nucleic acid encoding a murine CXc chemokine receptor.
 KW Mouse; CXc chemokine receptor; pre-B cell line DM34;
 KW CXc chemokine pre-B cell stimulatory factor PBSF/SDF-1;
 KW HIV infection; screening; inhibitor; AIDS; ds.
 OS Mus sp.
 PN M09835035-A1.
 PD 13-AUG-1998.
 PF 07-FEB-1997: J00299.
 PR 07-FEB-1997: WO-J00299.
 PA (SHIO) SHIONOGI & CO LTD.
 PI Iizasa H, Kishimoto T, Nagasawa T, Nakajima T, Tachibana K,
 PI Yoshida N, Yoshie O:
 WP: 98-447232/38.
 DR P-PSDB: M64778.
 DR Mouse CXc chemokine receptor binding to PBSF/SDF-1 pre-B cell
 PT stimulatory factor. Is useful for screening of potential HIV
 PT infection and AIDS inhibitors
 PT Claim 3: Pages 39-42; 76pp: Japanese.
 PS The present sequence encodes a murine CXc chemokine receptor which
 CC binds to the mouse CXc chemokine pre-B cell stimulatory factor
 CC PBSF/SDF-1. The nucleic acid is isolated from mouse pre-B cell
 CC line DM34. The receptor and cells expressing it can be used in the
 CC study and mapping of the mechanism of HIV infection and in screening
 CC of potential inhibitors of HIV infection and the development of AIDS.
 SO Sequence 1877 BP; 469 A; 451 C; 406 G; 551 T;

Query Match 2.9%; Score 50; DB 1; Length 1877;
 Best Local Similarity 100.0%; Pred. No. 1.2e-09;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1662 TAAACTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1711
 DB 1828 TAAACTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1877

RESULT 22

ID 052732 standard; DNA: 56 BP.
 AC 052732:
 DT 20-JUN-1994 (first entry)
 DE Sequence of oligo nucleotide adaptor used for the synthesis of
 DE asymmetrically tailed plasmid primers for use in cloning libraries.
 KW Asymmetrically tailed plasmid primer; adaptor; ds.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT misc-feature complement (1..4)
 FT /tag= a
 FT /label= Sticky end
 FT /tag= b
 FT /label= Sticky end
 PN US5270185-A.
 PD 14-DEC-1993.
 PF 21-APR-1989: 341523.
 PR 21-APR-1989: US-341523.
 PR 12-AUG-1992: US-928856.
 PA (HOFF) HOFFMANN LA ROCHE INC.
 PI Margolske RF:
 WP: 93-404926/50.
 DR Asymmetrically tailed plasmid primers - comprises cut,
 PT double-stranded DNA plasmid, useful for producing cloning
 PT libraries
 PS Claim 4: Figure 3A: 13pp: English.
 CC The inventors claim an asymmetrically tailed plasmid primer

CC comprising a *cat* gene, a functional origin of
CC replication and at least one functional selection marker gene, one 3'
CC terminus of which contains an oligo (dT) extension, the other 3'
CC terminus of which contains an oligo (dC) or oligo (dG) extension
CC terminated by a 3' phosphate group. More specifically, the
CC asymmetrically tailed plasmid primer is produced by SacI and KpnI
CC cleavage of plasmid pIV, and each terminus produced by the cleavage
CC is joined to the complementary adaptor, which is Q52731 or Q52732.
QO Sequence 56 bp: 1 A: 6 C: 8 G: 41 T:

Query Match	2.8%;	Score 49;	DB 1;	Length 56;
Best Local Similarity	100.0%;	Pred. NO. 5.8e-09;		
Matches 49;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	1674	AAAAAAAAAAAAAAAAAAAAAAAAAGCGGCGCC	1722
Db	56	AAAAAAAAAAAAAAAAAAAAAAAAAGCGGCGCC	8

Db 56 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCC 8

```

RESULT 23
052734/c
ID 052734 standard: DNA; 56 BP.
AC 052734:
DT 20-JUN-1994 (first entry)
DE Sequence of oligo nucleotide adaptor used for the synthesis of
DE asymmetrically tailed plasmid primers for use in cloning libraries
KM Asymmetrically tailed plasmid primer; adaptor; ds.
OS Synthetic.
FH FH Key Location/Qualifiers
FT FT misc-feature 1..4
FT FT *tag= a
FT FT /label= Sticky end
FT FT 16..56
FT FT misc-feature /tag= b
FT FT /label= Sticky end
PN US5270185-A.
PD 14-DEC-1993.
PE 21-APR-1989; 341523.
PR 21-APR-1989; US-341523.
PR 12-AUG-1992; US-928856.
PA (HOFF) HOFFMANN LA ROCHE INC.
PI Margolskee RF;
PI WPI: 93-404926/50.
PT Asymmetrically tailed plasmid primers - comprises cut,
PT double-stranded DNA plasmid, useful for producing cloning
PT libraries
PS Claim 5: Figure 3b: 13pp: English.
CC The inventors claim an asymmetrically tailed plasmid primer
CC comprising a cut, ds DNA plasmid contg. a functional origin of
CC replication and at least one functional selection marker gene, one
CC terminus of which contains an oligo (dT) extension, the other 3'
CC terminus of which contains an oligo (dC) or oligo (dG) extension
CC terminated by a 3' phosphate group. More specifically, the
CC plasmid primer is produced by EcoRI and HindIII cleavage
CC of plasmid pUC19, and each terminus produced by the cleavage
CC is joined to the complementary adaptor, which is Q52733 or Q52734.
Q Sequence 56 BP: 1 A; 5 C; 8 G; 42 T;

```

```
Query Match      2.8%; Score 49; DB 1; Length 56;
Best Local Similarity 100.0%; Pred. No. 5.8e-09;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0
```

Qy	1674	AAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCC	1722
Db	56	AAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCC	8

Db 56 AAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCC 8

RESULT	24
ID	T84975
AC	T84975; standard; cDNA; 2378 BP.

DT 27-Apr-1998 (first entry)
DE Rat glial cell derived neurotrophic factor receptor alpha cDNA.
KW Glial cell derived neurotrophic factor receptor alpha; GDNF alpha
KW GDNF; rat; kidney disease; glomerulonephritis; therapy; ds.
KW Rattus sp.

```
Query Match      2.8%; Score 49; DB 1; Length 2378;
Best Local Similarity 100.0%; Pred. No. 2.7e-09;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0
```

Qy	1665	ACTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	1713
Db	2250	ACTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	2298

Db 2250 ACTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2298

```

RESULT 25
V33197
AC V33197 standard; cDNA; 3443 bp.
ID V33197;
DT 21-DEC-1998 (first entry)
DE Sequenced protein DN747.7 cDNA.
KW Sequenced protein; DN747_7; human; ds
OS Homo sapiens.
FT FTCD Location/Qualifiers
FT CDS 1599..2546
FT /tag= a

```

```

PN MO6838209-A2.
PD 03-SEP-1998.
PF 25-FEB-1998; U03697.
PR 24-FEB-1998; US-028724.
PA (GEMV) GENETICS INST INC.
PI Agostino M, Jacobs K, Lavallie ER, McCoy JM, Merberg D,
PI Racie LA, Spaulding V, Treacy M;
PI WPI: 98-481139/41.
DR P-PSDB: W70325.
PT New isolated polynucleotide(s) and encoded polypeptide(s) -
PT obtained from human foetal kidney, adult colon, adult brain, foetal
PT brain and placenta cDNA libraries.
PS Claim 29: Page 75-77: 103pp: English.
CC This full-length cDNA clone, designated DN747_7, codes for a novel
CC human secreted protein (see W70325). It was isolated from a human
CC foetal brain cDNA library using a method selective for cDNAs
CC encoding secreted proteins, or was identified as encoding a
CC secreted or transmembrane protein on the basis of computer analysis
CC of the amino acid sequence of the encoded protein. Database
CC searching revealed some similarity between DN747_7 and some known
CC sequences. The invention provides new isolated polynucleotides
CC (see V33190-99), from human foetal kidney, adult colon, adult brain,
CC foetal brain and placenta cDNA libraries, that code for secreted
CC proteins (see W70319-27). They can be isolated from composite
CC deposit ATCC 98337 using specific probe sequences (see V33200-08).
CC The clones can be used for recombinant production of the
CC polypeptides, which may have activities such as e.g. nutritional
CC activity, cytokine and cell proliferation/differentiation activity,
CC immunostimulant or immunosuppressive, haematopoiesis regulating
CC activity, tissue growth activity, activin/inhibin activity,
CC chemotactic/chemokinetic activity, haemostatic and thrombolytic
CC activity, receptor/ligand activity, antiinflammatory activity,
CC cadherin/tumour invasion suppressor activity, tumour inhibition
CC activity, or other activities.
SQ Sequence 3443 BP; 875 A; 934 C; 834 G; 766 T;

Query Match 2.8%; Score 49; DB 1; Length 3443;
Best Local Similarity 100.0%; Pred. No. 2.5e-09;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1673 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCGC 1721
DB 1508 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCGC 1556

RESULT 26
V54124
ID V54124 standard; DNA; 2519 BP.
AC V54124:
DT 17-DEC-1998 (first entry)
DE Human membrane protein BA0306 coding sequence.
KW Membrane protein; BA0306; arteriosclerosis; coronary restenosis;
KW therapy: human; ds.
OS Homo sapiens.
FH Key
FT Location/Qualifiers
FT 1..1785
FT /*tag= a
FT /transl_except= (pos: 289..291, aa: Xaa)
FT /transl_except= (pos: 1015..1017, aa: Xaa)
FT /transl_except= (pos: 1492..1494, aa: Xaa)
FT /transl_except= (pos: 1540..1542, aa: Xaa)
FT /transl_except= (pos: 1582..1584, aa: Xaa)
FT /transl_except= (pos: 1609..1611, aa: Xaa)
FT /transl_except= (pos: 1693..1695, aa: Xaa)
FT /note= "Xaa= unspecified amino acid"

MO9838305-A1.
PD 03-SEP-1998.
PF 27-FEB-1998; J00835.
PR 25-FEB-1998; JP-062263.
PR 28-FEB-1997; JP-062259.
PA (NISB.) JAPAN TOBACCO INC.

```

```

PI Nakamura Y, Tanaka T, Tsukada S;
DR WPI: 98-481206/41.
DR P-PSDB: W74580.
PT Membrane protein(s) BA0306 and BA2303 - useful for, e.g. treatment
PT and prevention of arteriosclerosis and restenosis
PS Claim 16: Page 90-95: 141pp: Japanese.
CC This sequence encodes the human BA0306 membrane protein of the invention.
CC The invention also relates to the human BA2303 membrane protein. The two
CC membrane proteins are specifically expressed in mammals during
CC arteriosclerosis and coronary restenosis. The membrane proteins,
CC fragments of them, and antibodies against them are useful in the
CC treatment and prevention of arteriosclerosis and restenosis. Transgenic
CC mice expressing the extracellular region of the membrane proteins are
CC useful as models for studying these disorders.
SQ Sequence 2519 BP; 766 A; 439 C; 492 G; 809 T;

Query Match 2.8%; Score 49; DB 1; Length 2519;
Best Local Similarity 100.0%; Pred. No. 2.6e-09;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1673 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCGC 1721
DB 2434 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCGC 2482

RESULT 27
X52242
ID X52242 standard; DNA; 1770 BP.
AC X52242:
DT 25-JUN-1999 (first entry)
DE Protein PRO33 cDNA clone DNA34436-1238.
KW Secreted protein; transmembrane protein; human; enterocolitis;
KW Zollinger-Ellison syndrome; gastrointestinal ulceration;
KW congenital microvillus atrophy; skin disease; cell growth;
KW abnormal keratinocyte differentiation; psoriasis; epithelial cancer;
KW Parkinson's disease; Alzheimer's disease; ALS; neuropathy;
KW fibromodulin; dermal scarring; Usher Syndrome; Atrophia areata;
KW anti-thrombotic; wound healing; tissue repair; ss.
OS Homo sapiens.
PN MO9914328-A2.
PD 25-MAR-1999.
PF 16-SEP-1998; U19330.
PR 25-NOV-1997; US-066840.
PR 17-SEP-1997; US-059113.
PR 17-SEP-1997; US-059115.
PR 17-SEP-1997; US-059117.
PR 17-SEP-1997; US-059119.
PR 17-SEP-1997; US-058121.
PR 17-SEP-1997; US-058122.
PR 17-SEP-1997; US-059184.
PR 18-SEP-1997; US-059263.
PR 18-SEP-1997; US-059266.
PR 15-OCT-1997; US-062125.
PR 17-OCT-1997; US-062285.
PR 17-OCT-1997; US-062287.
PR 17-OCT-1997; US-062287.
PR 21-OCT-1997; US-063486.
PR 24-OCT-1997; US-062814.
PR 24-OCT-1997; US-062816.
PR 24-OCT-1997; US-063045.
PR 24-OCT-1997; US-063120.
PR 24-OCT-1997; US-063121.
PR 24-OCT-1997; US-063127.
PR 24-OCT-1997; US-063128.
PR 27-OCT-1997; US-063329.
PR 27-OCT-1997; US-063327.
PR 28-OCT-1997; US-063541.
PR 28-OCT-1997; US-063542.
PR 28-OCT-1997; US-063544.
PR 28-OCT-1997; US-063549.
PR 28-OCT-1997; US-063550.
PR 28-OCT-1997; US-063564.
PR 29-OCT-1997; US-063435.

```

29-OCT-1997: US-053704.
PR 29-OCT-1997: US-053732.
PR 29-OCT-1997: US-053738.
PR 29-OCT-1997: US-053734.
PR 29-OCT-1997: US-054215.
PR 29-OCT-1997: US-053735.
PR 31-OCT-1997: US-053870.
PR 31-OCT-1997: US-054103.
PR 03-NOV-1997: US-054248.
PR 07-NOV-1997: US-054809.
PR 12-NOV-1997: US-055186.
PR 17-NOV-1997: US-055846.
PR 18-NOV-1997: US-055693.
PR 21-NOV-1997: US-055120.
PR 21-NOV-1997: US-056354.
PR 24-NOV-1997: US-056772.
PR 24-NOV-1997: US-056466.
PR 24-NOV-1997: US-056770.
PR 24-NOV-1997: US-056511.
PR 24-NOV-1997: US-056453.
PA (GETH) GENENTECH INC.
PI Chen J, Goddard A, Gurney AL, Pennica D, Wood WI, Yuan J:
DR WPI: 99-229533/19.
P-PSDB: Y13371.
PT New isolated human genes and polypeptides used in, e.g. treatment of
PS gastrointestinal ulceration
PS Claim 2: Fig 57: 320pp: English.
CC X52213-74 encode secreted and transmembrane human proteins, and are
CC obtained from cDNA libraries, prepared from fetal lung, fetal kidney,
CC fetal brain, fetal liver and fetal retina. The encoded polypeptides
CC have specific uses based on their homology to known polypeptides,
CC e.g. PRO211 and PRO217 can be used for disorders associated with the
CC preservation and maintenance of gastrointestinal mucosa and the repair
CC of acute and chronic mucosal lesions (e.g. enterocolitis,
CC Zollinger-Ellison syndrome, gastrointestinal ulceration and congenital
CC microvillus atrophy), skin diseases associated with abnormal keratinocyte
CC differentiation (e.g. psoriasis, epithelial cancers such as lung squamous
CC cell carcinoma of the vulva and gliomas), potent effects on cell growth
CC and development, diseases related to growth or survival of nerve cells
CC including Parkinson's disease, Alzheimer's disease, ALS, neuropathies or
CC cancer. PRO265 can be used as for fibromodulin, e.g. for reducing dermal
CC scarring. PRO264 can be used as a target for anti-tumor drugs. PRO333
CC may be used in the treatment of Usher Syndrome or Atrophila areata;
CC PRO269 can be used as an anti-thrombotic agent; PRO287 polypeptides
CC and portions may have therapeutic applications in wound healing and
CC tissue repair; PRO317 can be used for treating problems of the kidney,
CC uterus, endometrium, blood vessels, or related tissue, e.g. in the
CC heart of genital tract. 621 A: 313 C: 329 G: 507 T:
SQ Sequence 1770 BP: 621 A: 313 C: 329 G: 507 T:

Query Match 2.8%; Score 49; DB 1; Length 1770;
Best Local Similarity 100.0%; Pred. No. 2.8e-09;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1665 ACTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1713
DB 1663 ACTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1711

FT W09318158-A. /note- "Has acyl-ACP thioester activity."
PD 16-SEP-1993.
PF 03-MAR-1993; G00432.
PR 03-MAR-1992; GB-004583.
PA (UNIL) UNILEVER NV.
PA (UNIL) UNILEVER PLC.
PI Hellyer SA, Loader NM, Safford R, Slabas AR:
DR WPI: 93-303472/38.
P-PSDB: R41674.
PT CDNA encoding enzyme with acyl-ACP-thioesterase activity -
PT useful for transforming plants, for altering properties e.g. seed
PT storage oil composition
PS Claim 2: Figure 2: 48pp: English.
CC The nucleotide sequence is useful for transforming plant cells and
CC altering that plants characteristics. Characteristics which may be
CC altered are (1) altering the level of acyl-ACP thioesterase
CC activity; (2) fatty acid yield and/or fatty acid composition and (3)
CC yield and/or composition of seed storage oil.
SQ Sequence 1586 BP: 491 A: 324 C: 339 G: 432 T:

Query Match 2.8%; Score 48; DB 1; Length 1586;
Best Local Similarity 100.0%; Pred. No. 5.6e-09;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1666 CTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1713
DB 1521 CTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1568

RESULT 29
058663
ID O58663 standard: cDNA to mRNA; 699 BP.
AC O58663;
DT 14-OCT-1994 (first entry)
DE Human superoxide dismutase gene.
KW Superoxide dismutase; SOD; transformation: production: ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT cds 68..529
FT /tag= a
FT /product= superoxide_dismutase
FT 625..699
FT /tag= b
PN poly-a-signal
PD J06046860-A.
PD 22-FEB-1984.
PF 23-SEP-1992; 279193.
PR 23-SEP-1992; JP-279193.
PA (NIPK) NIPPON KAYAKU KK.
DR WPI: 94-097023/12.
DR P-PSDB: R47062.
PT DNA encoding human super:oxide dismutase (SOD) and transformed
PT microorganism - useful for prodn. of SOD
PS Claim 1: Page 7: 10pp: Japanese
CC The sequence encodes a superoxide dismutase. The gene may be used to
CC produce superoxide dismutase in increased yields by transforming it
CC into E.coli.
SQ Sequence 699 BP: 242 A: 127 C: 181 G: 149 T:

Query Match 2.8%; Score 48; DB 1; Length 699;
Best Local Similarity 100.0%; Pred. No. 7.8e-09;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1666 CTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1713
DB 622 CTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 669

RESULT 30
075361
ID Q75361 standard: cDNA to mRNA; 1490 BP.

```

AC 075361:
DT 01-JUL-1995 (first entry)
DE Gibberellin-20-oxidase.
KW Gibberellin-20-oxidase; plasmid PAT2353; transgenic plant;
  crop improvement; ss.
OS Arabidopsis thaliana.
FH key Location/Qualifiers
FT cds 98..1231
FT /*tag= a
FT /*note= "GA-20-oxidase"
PN MO9428141-A.
PD 08-DEC-1994.
PF 24-MAY-1993; GB-011147.
PR 28-MAY-1993; GB-011147.
PA (LONG-) LONG ASHTON RES STATION.
PI Graebe JE, Hedden P, Lange T, Phillips A;
DR P-PSDB: R64144.
PT New DNA encoding gibberellin-20-oxidase and related vectors -
PT also host cells and transgenic plants, used to modulate plant
PT growth and development, e.g. anti-sense constructs will reduce
PT vegetative growth.
PS Claim 10: Page 78-79; 97pp: English.
CC Clone PAT353 containing DNA encoding gibberellin-20-oxidase may be
CC used to construct chimeric genes and binary vectors to transform
CC plants by usual methods, e.g. Agrobacterium-mediated transfer.
CC propagated transgenic plants show modulated growth and development
CC e.g. reducing GA-20-oxidase expression with antisense constructs
CC reduces vegetative growth in cereals and grasses, improving
CC resistance to frost, etc. Sense constructs improve fruit set and
CC growth, extended stems and leaves, induce flowering and overcome
CC dormancy, etc. Preferred transgenic plants are maize and wheat.
SQ Sequence 1490 BP; 486 A; 330 C; 274 G; 400 T;

Query Match 2.8%; Score 48; DB 1; Length 1490;
Best Local Similarity 100.0%; Pred. No. 6.7e-09;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1666 CTTAAAAA... 1713
DB 1421 CTTAAAAA... 1468

RESULT 31
Q99552 standard; cDNA to mRNA; 861 BP.
AC Q99552:
DT 29-APR-1996 (first entry)
DE Thrombopoietin coding sequence.
KW Thrombopoietin; TPO; humoral factor; platelet; antibody; therapy; AIDS;
  thrombocytopenia; hypoplastic anaemia; thrombotic thrombocytopenia;
  disseminated intravascular coagulation syndrome; ds.
OS Homo sapiens.
FH key Location/Qualifiers
FT cds 25..783
FT /*tag= a
FT /*product= thrombopoietin
FT signal_peptide 25..87
FT /*tag= b
FT mat_peptide 88..783
FT /*tag= c
PN MO9521919-A2.
PD 17-AUG-1995.
PF 14-FEB-1995; J00208.
PR 14-FEB-1994; JP-039090.
PR 14-MAR-1994; US-212164.
PR 25-MAR-1994; JP-079842.
PR 01-APR-1994; US-221020.
PR 01-JUN-1994; JP-155126.
PR 15-JUN-1994; JP-167328.
PR 20-JUL-1994; US-278083.
PR 17-AUG-1994; JP-193169.

```

```

PR 17-AUG-1994; JP-227159.
PR 18-AUG-1994; JP-193916.
PR 11-OCT-1994; US-320300.
PR 01-NOV-1994; JP-304167.
PR 01-DEC-1994; JP-298669.
PR 22-DEC-1994; US-361811.
PR 28-DEC-1994; JP-341200.
PR 31-JAN-1995; US-381478.
PA (KIRI) KIRIN BREWERY KK.
PI Akahori H, Iwamatsu A, Kato T, Kuroki R, Miyazaki H;
PI Muto T, Ohgami K, Shimizu T;
DR WPI: 95-293120/38.
DR P-PSDB: R81378.
PT Thrombopoietin polypeptide which specifically stimulates or
PT increases platelet production - useful in treatment of
PT thrombocytopenia, also related DNA and vectors
PS Claim 19; Page 337-338; 383pp: English.
CC This sequence represents the coding sequence for a thrombopoietin (TPO)
CC protein. TPO is a humoral factor capable of promoting platelet
CC production. The encoded protein can be purified by using a gel affinity
CC column where Mpl has been coupled to a resin. Q99551 and Q99553 also
CC encode TPO proteins. The DNA sequences are inserted into vectors which
CC are used to transform prokaryotic and eukaryotic host cells. Using the
CC proteins encoded by these sequences, and derivatives of them, antibodies
CC specifically immunoreactive with a TPO protein can be created. The
CC antibodies, DNA sequences and vectors are used to isolate the protein
CC sequences. The TPO proteins can then be used in the treatment of
CC platelet disorders. These include thrombocytopenia, hypoplastic anaemia,
CC AIDS, disseminated intravascular coagulation syndrome and thrombotic
CC thrombocytopenia.
SQ Sequence 861 BP; 243 A; 257 C; 197 G; 164 T;

Query Match 2.8%; Score 48; DB 1; Length 861;
Best Local Similarity 100.0%; Pred. No. 7.5e-09;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1666 CTTAAAAA... 1713
DB 783 CTTAAAAA... 830

RESULT 32
T90174 standard; cDNA; 1474 BP.
AC T90174:
DT 20-APR-1998 (first entry)
DE Oil seed rape cysteine protease cDNA clone CDCY566.
KW Promoter; oil seed rape; cDNA clone CDCY566;
  cysteine protease; disrupter protein; plant; tolerance; herbicide;
  insect pest; fungal disease; improved yield; improved quality;
  fertility control; ss.
OS Brassica napus.
FH key Location/Qualifiers
FT cds 33..1157
FT /*tag= a
PN MO9735983-A2.
PD 02-OCT-1997.
PF 18-MAR-1997; G00729.
PR 22-MAR-1996; GB-006062.
PA (ZENE) ZENECA LTD.
PI Greenland AT, Jepson I, Thomas DRP;
DR WPI: 97-489646/45.
DR P-PSDB: W27446.
PT Cysteine protease promoter - obtained from oil seed rape, useful to
PT develop plants with improved agronomic characteristics
PS Claim 23; Fig 17; 137pp: English.
CC The present sequence encodes an oil seed rape cysteine protease,
CC the promoter for which can restrict the expression of a disrupter
CC protein gene to a suitable stage of plant development, to provide
CC plants with novel agronomic features, e.g. tolerance to herbicides,
CC insect pests and fungal diseases, improved yield and/or quality of
CC harvested product, and novel mechanisms for the control of plant

```


CC fertility. 1474 BP; 528 A; 265 C; 316 G; 365 T;
 SO Sequence

Query Match 2.8%; Score 48; DB 1; Length 1474;
 Best Local Similarity 100.0%; Pred. No. 6,7e-09;
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1666 CTTAAAAA
 DB 1373 CTTAAAAA

RESULT 33

V21646 V21646 standard; cDNA; 2312 BP.

AC V21646:
 DT 21-JUL-1998 (first entry)
 DE T. gondii chorismate synthase encoding cDNA.
 KW Chorismate synthase; enzyme; Apicomplexan parasite; diagnosis;
 KM metabolic pathway; parasitic disease; psda gene; toxoplasma gondii;
 OS Toxoplasma gondii.
 FH Key Location/Qualifiers
 FT CDS 152..1772
 FT /tag= a
 FT /product= "chorismate synthase enzyme"

PT W0980361-A2.
 PD 29-JAN-1998.
 PF 18-JUL-1997; U12497.
 PR 13-JUN-1997; US-049620.
 PR 19-JUL-1996; US-022209.
 PR 23-DEC-1996; US-773302.
 PR 17-MAR-1997; US-040849.
 PA (ARCH-) ARCH DEV CORP.
 PI Johnson JJ, McLeod RLM, Mels L, Roberts CW, Roberts F;
 DR WPI: 98-120782/11.
 DR P-PSDB: W53105.
 PT Use of Apicomplexan parasite components - from plant-like metabolic
 PT pathways for developing products to treat, diagnose or prevent

PT Parasitic diseases
 PS Claim 12: Pages 172-175, 211pp; English.
 CC This cDNA encodes a T. gondii chorismate synthase enzyme. A component of
 CC a plant-like metabolic pathway in an Apicomplexan parasite used for
 CC producing a composition that interferes with the growth or survival of
 CC the Apicomplexan parasite is selected from this 2312 base pair T. gondii
 CC chorismate synthase enzyme sequence or a 1837 base pair T. gondii
 CC chorismate synthase enzyme sequence. The pathway does not involve the
 CC psda gene or PPI phosphotransferase and the component is not encoded by
 CC the plasmid genome, and is not generally operative in animals. The
 CC component comprises a transit peptide when the pathway involves transport
 CC of proteins into or out of an organelle. An Apicomplexan parasite
 CC comprising a gene encoding a component of a metabolic pathway which
 CC operates for the growth and survival of the parasite, where the gene is
 CC altered can be used to construct a vaccine for protecting an animal
 CC against infection by the parasite. The invention provides methods for
 CC identifying such components and for developing antibodies and inhibitors
 CC to the component. The products and methods can be used to develop
 CC products to treat, diagnose and prevent diseases due to Apicomplexan
 CC parasites such as Toxoplasma gondii causing toxoplasmosis. Plasmodia
 CC causing malaria, Cryptosporidia causing cryptosporidiosis, Eimeria
 CC causing eimeriosis, Babesia causing babesiosis, Theileria causing
 CC theileriosis and Neospora caninum.
 SO Sequence 2312 BP; 548 A; 640 C; 624 G; 500 T;

Query Match 2.8%; Score 48; DB 1; Length 2312;
 Best Local Similarity 100.0%; Pred. No. 6,1e-09;
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1666 CTTAAAAA
 DB 2249 CTTAAAAA

RESULT 34

V18187 V18187 standard; cDNA to mRNA; 5503 BP.

AC V18187:
 DT 28-AUG-1998 (first entry)
 DE Fanconi anaemia of complementation group A gene.
 KW Fanconi anaemia of complementation group A; FA-A; genetic defect;
 KM prenatal FA-A; FA-A carrier detection; disease diagnosis; ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 32..4399
 FT /tag= a
 FT /product= FA-A

PN W09814462-A1.
 PD 09-APR-1998.
 PF 03-OCT-1997; U18010.
 PR 04-OCT-1996; US-726012.
 PA (FANC-) FANCONI ANEMIA RES FUND INC.
 PI Joenje H, Lo Ten Foe JR;
 DR WPI: 98-240012/21.
 DR P-PSDB: W48663.
 PT DNA for Fanconi Anaemia complementation group A - useful for, e.g.,
 PT developing products for diagnosis and screening of disease and gene
 PT therapy
 PS Claim 4; Fig 2; 63pp; English.
 CC This sequence encodes the Fanconi anaemia of complementation group A
 CC (FA-A) protein of the invention. The DNA's may be used to complement a
 CC genetic defect in a cell (especially the FA-A gene). The products can be
 CC used for screening (especially prenatal FA-A), detection of FA-A carriers
 CC and FA-A disease diagnosis
 SO Sequence 5503 BP; 1208 A; 1527 C; 1492 G; 1276 T;

Query Match 2.8%; Score 48; DB 1; Length 5503;
 Best Local Similarity 100.0%; Pred. No. 5,1e-09;
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1666 CTTAAAAA
 DB 5446 CTTAAAAA

RESULT 35

V37789/C V37789 standard; DNA; 90 BP.

AC V37789:
 DT 09-SEP-1998 (first entry)
 DE Analytical solid phase nucleic acid detection hybridisation probe 16.
 KW Analytical solid phase: detection; hybridisation; probe; target;
 KW diagnosis; screening; disease; drug sensitivity; transplantation organ;
 KW food industry; food poisoning; ss.
 OS Synthetic.
 PN W09811210-A1.
 PD 19-MAR-1998.
 PF 12-SEP-1997; J03232.
 PR 13-SEP-1996; JP-243720.
 PA (MOLE-) LAB MOLECULAR BIOPHOTONICS.
 PI Abe S, Sato Y;
 DR WPI: 98-271664/24.
 PT Analytical solid phase for detecting nucleic acids - contains a base
 PT sequence which hybridises with polynucleotide sequence of the
 PT target, and a set of probes immobilised on the solid phase via a
 PT linker

PS Example 7; Page 21-22; 37pp; Japanese.
 CC The present sequence represents a probe used in an example from the
 CC present invention. The present invention describes an analytical solid
 CC phase method for detecting nucleic acids. The method comprises obtaining
 CC a base sequence which hybridises with the polynucleotide sequence of the
 CC target, and a set of probes immobilised on the solid phase via a linker
 CC which is enzymatically ligated during hybridisation. The product is used
 CC for the detection of specific nucleic acids. Possible fields of

CC application include diagnosis of disease, detection of drug sensitivity,
 CC screening for appropriate transplantation organs, testing in the food
 CC industry to prevent food poisoning. A simple, quick method for
 CC selectively detecting target in a mixed sample is obtained.
 SO Sequence 90 BP: 6 A; 14 C; 13 G; 57 T;

Query Match 2.8%; Score 48; DB 1; Length 90;
 Best Local Similarity 100.0%; Pred. No. 1.2e-08;
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1665 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACG 1716
 ||||||||||||||||||||||||||||||||||||||||
 DB 65 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACG 18

RESULT 36
 V43610 36
 ID V43610 standard; DNA; 1049 BP.
 AC V43610;
 DT 24-SEP-1998 (first entry)
 DE Human secreted protein 10 encoding DNA.
 KW Secreted protein; human; cell proliferation; cytokine activity;
 KW tissue growth; cellular differentiation; regeneration; activin;
 KW inhibin; chemotactic; haemostatic; thrombolytic; tumour inhibition;
 KW anti-inflammatory activity; biomarker; ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 21..821
 FT /tag= a
 FT /product= "human secreted protein"

W09825959-A2.
 PD 18-JUN-1998.
 PF 11-DEC-1997; U22787.
 PR 11-DEC-1996; US-032757.
 PA (CHIR) CHIRON CORP.
 PI Escobedo J, Garcia F, Hu Q, Kothakota S, Williams LT;
 DR WPI: 98-348453/30.
 P-PSDB: W63690.
 PT Secreted human polypeptides - having cytokine, cell proliferation or
 PT differentiation, activin or inhibin, tumour inhibition or
 PT anti-inflammatory activities
 PS Claim 6: Page 37; 78pp; English.
 CC This DNA encodes a human secreted protein. The specification provides
 CC secreted protein sequences (W63681 to W63699) encoded by the nucleic
 CC acid sequences shown in V43601 to V43619. The invention provides a
 CC method of identifying a secreted polypeptide which is modified by rough
 CC microsomes. The secreted proteins can be used in assays to determine
 CC biological activities, such as cytokine, cell proliferation, or cellular
 CC differentiation activities, tissue growth or regeneration, activin or
 CC inhibin activity, chemotactic or chemokinetic activity, haemostatic or
 CC thrombolytic activity, receptor/ligand activity, tumour inhibition, or
 CC anti-inflammatory activity. The proteins can also be used as biomarkers,
 CC to identify tissues or cell types which express the proteins, or a stage-
 CC or disease-specific alteration in protein expression. They can be used
 CC in protein interaction assays, to identify ligands or binding proteins.
 CC Compounds which affect the biological activities of the secreted proteins
 CC or their ability to interact with specific ligands can be identified
 CC using the proteins in screening assays. The proteins and antibodies that
 CC bind specifically to the protein can also be used to design diagnostic
 CC tests and therapeutic compositions for diseases which may be associated
 CC with altered expression of these proteins. Fusion proteins comprising,
 CC e.g. signal sequences or transmembrane domains of the proteins can be
 CC used to target other protein domains to cellular membrane or they can
 CC be secreted extracellularly.
 SO Sequence 1049 BP: 345 A; 214 C; 206 G; 284 T;

Query Match 2.8%; Score 48; DB 1; Length 1049;
 Best Local Similarity 100.0%; Pred. No. 7.2e-09;
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1665 CTTAAAAA.....1713

DB 981 CTTAAAAA.....1028

RESULT 37
 V58363 37
 ID V58363 standard; cDNA; 772 BP.
 AC V58363;
 DT 25-NOV-1998 (first entry)
 DE Coding sequence for secreted protein of clone EQ219.1.
 KW Secreted protein; nutritional activity; immune stimulant;
 KW immune suppressor; haematopoiesis regulator; tissue growth activity;
 KW activin/inhibin activity; chemotactic/chemokinetic activity;
 KW haemostatic activity; thrombolytic activity; receptor/ligand activity;
 KW anti-inflammatory activity; cadherin suppressor; tumour inhibitor;
 KW tumour invasion suppressor; therapy; ds.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 262..654
 FT /tag= a
 FT /note= "no stop codon given"

W09837094-A2.
 PD 27-AUG-1998.
 PF 24-FEB-1998; U03595.
 PR 23-FEB-1998; US-028168.
 PR 24-FEB-1997; US-804561.
 PA (GENY) GENENTEC INST INC.
 PI Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D,
 PI Racle LA, Spaulding V, Treacy M;
 DR WPI: 98-467491/40.
 P-PSDB: W69347.
 PT Secreted proteins with biological activity - and encoding
 PT poly(nucleotide)s, useful e.g. in therapy and diagnosis of medical
 PT conditions and to identify protein agonists or antagonists
 PS Claim 35: Page 86-87; 11pp; English.
 CC This sequence encodes a secreted protein of the invention. This sequence
 CC is referred to as clone EQ219.1, and was isolated from a human adult
 CC testes cDNA library. The polypeptides are predicted to have useful
 CC biological activities which would make them suitable for treating
 CC preventing or ameliorating medical conditions in humans and animals,
 CC although no supporting biological data is given. Suggested activities,
 CC include nutritional activity, immune stimulating or suppressing activity,
 CC haematopoiesis regulating activity, tissue growth activity,
 CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory
 CC activity, cadherin/tumour invasion suppressor activity, tumour inhibition
 CC activity or other activities. They may also be used for diagnostic
 CC purposes. The polynucleotides are useful to produce the proteins or
 CC generate probes or primers to identify and/or amplify similar genes
 CC e.g. species homologues. They are also useful for gene therapy and to
 CC produce transgenic animals with altered gene expression.
 SO Sequence 772 BP: 301 A; 164 C; 132 G; 175 T;

Query Match 2.8%; Score 48; DB 1; Length 772;
 Best Local Similarity 100.0%; Pred. No. 7.6e-09;
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1665 CTTAAAAA.....1713
 ||||||||||||||||||||||||||||||||||||||||
 DB 651 CTTAAAAA.....698

RESULT 38
 V58754 38
 ID V58754 standard; cDNA; 1328 BP.
 AC V58754;
 DT 18-JAN-1999 (first entry)
 DE Human secreted protein ax318.3 cDNA.
 KW Secreted protein; human; ax318.3; ds.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 521..1114

```

FT      sig_peptide      /*tag= a
FT      563..601
FT      /*tag= b
FT      mat_peptide      602..1111
FT      /*tag= c
PN      WO980404-A2.
PD      17-SEP-1998.
PR      09-MAR-1998: U04601.
PR      06-MAR-1998: US-016321.
PR      11-MAR-1997: US-815381.
PA      (GEMV ) GENETICS INST INC.
PI      Agostino M, Jacobs K, Lavallie ER, McCoy JM, Merberg D,
PI      Racie LA, Spaulding V, Treacy M.
PI      WPI: 98-520802/44.
DR      P-PSDB: W69423.
PT      New isolated polynucleotides and secreted proteins - useful as, e.g.
PT      nutritional additives, immunostimulators, haemotopoiesis regulators
PT      and as diagnostic agents
PS      Claim 1: Page 56: 11pp: English.
CC      This full-length cDNA clone, designated ax318.3, codes for a novel
CC      human secreted protein (see W69423), i.e. ax318.3 protein. It was
CC      isolated from a human adult testis cDNA library using methods which
CC      are selective for cDNAs encoding secreted proteins, or was
CC      identified as encoding a secreted or transmembrane protein on the
CC      basis of computer analysis of the encoding protein. Homology is
CC      shown to some database sequences. The invention provides isolated
CC      polynucleotides (see V58754-63) obtained from human adult testis,
CC      brain, retina or placenta, or from foetal kidney or brain cDNA
CC      libraries. These are all deposited as ATCC 98353. They encode
CC      novel human secreted proteins (see W69423-33) that may have e.g.
CC      nutritional activity, immune stimulating or suppressing activity,
CC      haemotopoiesis regulating activity, tissue growth activity,
CC      activin/inhibin activity, chemotactic/chemokinetic activity,
CC      haemostatic and thrombolytic activity, receptor/ligand activity,
CC      antiinflammatory activity, cadherin/tumour invasion suppressor
CC      activity, tumour inhibition or other activities. They may also be
CC      used for diagnostic purposes.
SQ      Sequence 1328 BP: 446 A: 307 C: 282 G: 293 T:

```

```

Query Match      2.8%; Score 48; DB 1: Length 1328;
Best Local Similarity 100.0%; Pred. No. 6.8e-09;
Matches 48: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1659 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCG 1716
DB      452 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCG 499

```

```

RESULT 39
ID      V52969 standard: cDNA: 1373 BP.
AC      V52969:
DT      11-JAN-1999 (first entry)
DE      Human G1 protein Isoform beta (CASH-beta) cDNA.
KW      G1 protein; CASH-beta; human; caspase homologue; Fas receptor;
KW      modular; apoptosis; cell death; inflammation; tumour; HIV;
KW      therapy; ss.
OS      Homo sapiens.
FH      key      Location/Qualifiers
FT      CDS      482..1137
FT      /*tag= a
PN      WO9839435-A1.
PD      11-SEP-1998.
PD      26-FEB-1998: IL0098.
PR      01-MAY-1997: IL-120759.
PR      03-MAR-1997: IL-120367.
PA      (YEDA ) YEDA RES & DEV CO LTD.
PI      Brodalski V, Goltsev Y, Kovaleenko A, Varfolomeev E,
PI      Wallach D:
PI      WPI: 98-495842/42.
DR      P-PSDB: W78904.
PT      New DNA encoding isoforms of G1 protein which bind MORT-1 - and

```

```

PT      regulate the effects of FAS and tumour necrosis factor receptors,
PT      useful for killing of cells e.g. HIV and cancer cells
PS      Claim 4: Fig 2: 13pp: English.
CC      This cDNA sequence codes for the beta isoform (see W78904) of novel
CC      human G1 protein. It was isolated from a human skin fibroblast
CC      cDNA library using a cDNA probe corresponding to the G1 sequence.
CC      G1-beta (also called CASH beta, CASH being caspase homologue)
CC      and a longer isoform, G1-alpha (see W78903), represent 2 splice
CC      variants of the G1 protein. These G1 proteins are capable of
CC      binding to, or interacting directly or indirectly, with MORT-1
CC      or with MORT-binding proteins such as Mch4 (CASP-10) and MCH4
CC      (CASP-8), and thereby of binding to the intracellular domain of
CC      the FAS-R receptor, to which MORT-1 binds, or of binding to the
CC      intracellular domain of the p55 tumour necrosis factor (TNF)
CC      receptor, to which TRADD binds and to which TRADD protein MORT-
CC      1 binds. Hence, they are considered as mediators or modulators of
CC      FAS-R having a role in e.g. the signalling process that is initiated
CC      by the binding of FAS ligand to FAS-R, and also having a role in
CC      the signalling process initiated by the binding of TNF to p55-R.
CC      G1 DNA (I) and polypeptide (II), vectors and fragments are used to
CC      regulate cell death or inflammatory processes. (II) is used to
CC      inhibit cell death, and its inhibitors augment/enhance the
CC      processes. (I) and (II) regulate the FAS-R ligand or TNF effect on
CC      cells carrying an FAS-R or p55-R. Tumour, HIV-infected or other
CC      diseased cells can be treated using a viral vector encoding a viral
CC      surface protein, which binds a specific cell surface receptor and a
CC      sequence encoding (II), which kills the cell. Antisense
CC      oligonucleotides, introduced using the above vector, block the
CC      expression of (II) and can also regulate the above effects. These
CC      effects can also be regulated using a vector encoding a ribozyme
CC      that interacts with a cellular mRNA encoding (II), and allows (II)
CC      expression.
SQ      Sequence 1373 BP: 417 A: 283 C: 335 G: 338 T:

```

```

Query Match      2.8%; Score 48; DB 1: Length 1373;
Best Local Similarity 100.0%; Pred. No. 6.8e-09;
Matches 48: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1656 CTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1713
DB      1310 CTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1357

```

```

RESULT 40
ID      V64590 standard: DNA: 1579 BP.
AC      V64590:
DT      29-JAN-1999 (first entry)
DE      Tobacco chili DNA.
KW      ChlD: magnesium chelatase subunit D; tobacco; transgenic plant; activity;
KW      subunit I; subunit H; effector; detectable; ss.
OS      Nicotiana tabacum.
PN      DE19717656-A1.
PD      29-OCT-1998.
PE      25-APR-1997: 017656.
PR      25-APR-1997: DE-017656.
PA      (AGRE ) HOECHST-SCHERING AGREVO GMBH.
PI      Graefe S, Grimm B, Haendel F, Papenbrock J, Schmidt F,
PI      Streiber W:
PI      WPI: 98-569720/49.
DR      New nucleic acid encoding plant magnesium chelatase subunit D
PT      protein - used to amplify nucleic acid, produce antibodies and
PT      transgenic plants and to identify plant magnesium chelatase
PT      effectors
PS      Example 6: Page 23-24: 30pp: German.
CC      This sequence encodes a tobacco magnesium chelatase subunit I (chl1)
CC      which is used in a method to amplify, isolate or identify a nucleic
CC      acid molecule coding for a chlD protein or a biologically active
CC      fragment, to produce antibodies and to produce transgenic plants. The
CC      chlD protein is used for magnesium chelatase activity determination.
CC      Determination of the interaction of plant magnesium chelatase subunits
CC      comprises transforming a host cell with the nucleic acid and with at

```

CC least one DNA sequence encoding another magnesium chelataase subunit, in
 CC such a way that the interaction of the magnesium chelataase subunits
 CC produces a detectable signal, preferably by activation of a reporter
 CC gene. Determination of the activity of a plant magnesium chelataase
 CC comprises contacting the CHD protein with the gene products of DNA
 CC sequences coding for magnesium chelataase subunits I and H in such a
 CC way that the enzymatic activity of the magnesium chelataase subunits
 CC produces a detectable signal. The above methods are used to identify
 CC plant magnesium chelataase effectors.
 CC Sequence 1579 BP; 490 A; 285 C; 368 G; 436 T;

Query Match 2.8%; Score 48; DB 1; Length 1579;
 Best Local Similarity 100.0%; Pred. No. 6.6e-09;
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1666 CTTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1713
 DB 1508 CTTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1555

RESULT 41

X33812
 ID X33812 standard; DNA; 2120 BP.
 AC X33812.
 DT 25-JUN-1999 (first entry)
 DE Coding sequence for human secreted protein cb96.10.
 KW Secreted protein: human; nutritional activity; cytokine; tissue growth;
 KW cell proliferation; cell differentiation; immune stimulant; chemotaxis;
 KW immune suppressant; haematopoiesis regulator; activin; inhibin; cadherin;
 KW chemokinesis; haemostasis; thrombolysis; anti-inflammatory; gene therapy;
 KW tumour invasion suppressor; tumour inhibitor; ss.
 OS Homo sapiens.
 PN WO9913066-A1.
 PD 13-MAR-1999.
 PF 03-SEP-1998; U18724.
 PR 03-SEP-1997; US-929007.
 PA (GENY) GENETICS INST INC.
 PI Agostino MJ, Evans C, Jacobs K, Lavalie ER, McCoy JM,
 PI Merberg D, Racie LA, Spaulding V, Treacy M;
 PI WPI: 98-229235/19.
 DR P-PSDB: Y05319.
 PT New polynucleotides encoding secreted human proteins
 PS Claim 17; Page 80-81; 96pp; English.
 CC This sequence encodes a human protein of the invention. The
 CC secreted proteins were obtained from human adult placenta, foetal brain,
 CC adult testes or adult blood cDNA libraries. The polynucleotides (PNS) and
 CC proteins are predicted to have biological activities which would make
 CC them suitable for treating, preventing or ameliorating medical conditions
 CC in humans and animals, although no supporting data is given. Suggested
 CC activities include nutritional activity, cytokine and cell
 CC proliferation/differentiation activity, immune stimulating
 CC (e.g. as vaccines) or suppressing activity, haematopoiesis regulating
 CC activity, tissue growth activity, activin/inhibin activity,
 CC chemotactic/chemokinetic activity, haemostatic and thrombolytic activity,
 CC receptor/ligand activity, anti-inflammatory activity, cadherin/tumour
 CC invasion suppressor activity, and tumour inhibition activity. The PNS are
 CC also stated to be useful for gene therapy.
 CC Sequence 2120 BP; 667 A; 347 C; 403 G; 703 T;

Query Match 2.8%; Score 48; DB 1; Length 2120;
 Best Local Similarity 100.0%; Pred. No. 6.2e-09;
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1666 CTTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1713
 DB 2042 CTTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2089

RESULT 42

X37521
 ID X37521 standard; cDNA; 1090 BP.

AC X37521; (first entry)
 DT 06-JUL-1999
 DE Human secreted protein cDNA fragment containing gene 71.
 KW Human; secreted protein; treatment; prevention; protein therapy; AIDS;
 KW gene therapy; diagnosis; cancer; tumour; neurodegenerative disorder;
 KW developmental abnormality; fetal deficiency; blood disorder; leukemia;
 KW immune system disease; autoimmune disease; hepatic disease; lymphoma;
 KW renal disease; inflammation; allergy; Alzheimer's disease; schizophrenia;
 KW cognitive disorder; prostate disease; skeletal; cardiac; muscle disorder;
 KW pulmonary disorder; transplant rejection; osteoclast; osteoporosis;
 KW arthritis; malignancy; digestive; endocrine; infection; ss.
 OS Homo sapiens.
 PN WO9918208-A1.
 PD 15-APR-1999.
 PF 01-OCT-1998; U20775.
 PR 02-OCT-1997; US-060884.
 PR 02-OCT-1997; US-060833.
 PR 02-OCT-1997; US-060836.
 PR 02-OCT-1997; US-060837.
 PR 02-OCT-1997; US-060838.
 PR 02-OCT-1997; US-060839.
 PR 02-OCT-1997; US-060843.
 PR 02-OCT-1997; US-060862.
 PR 02-OCT-1997; US-060866.
 PR 02-OCT-1997; US-060874.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Carrier KC, Duan DR, Endress GA, Peng P, Ferrie AM,
 PI Florence KA, Greene JM, Janat F, Lafleur DW, Ni J,
 PI Rosen CA, Ruben SM, Shi Y, Young P, Yu G;
 PI WPI: 99-264022/22.
 DR P-PSDB: Y07922.

PT New isolated human genes and the secreted polypeptides they encode
 PS Claim 1a; Page 263; 368pp; English.
 CC This invention describes novel isolated human genes and the secreted
 CC proteins they encode. The products of the invention are useful for
 CC preventing, treating or ameliorating medical conditions, e.g. by protein
 CC or gene therapy. Also pathological conditions can be diagnosed by
 CC determining the amount of the new polypeptides in a sample or by
 CC determining the presence of mutations in the new polynucleotides.
 CC Specific uses are described for each of the 101 polynucleotides, based on
 CC which tissues they are most highly expressed in, and include developing
 CC products for the diagnosis or treatment of cancer, tumours,
 CC neurodegenerative disorders, developmental abnormalities and fetal
 CC deficiencies, blood disorders, leukemias, diseases of the immune system,
 CC autoimmune diseases, hepatic and renal disease, lymphomas, inflammation,
 CC allergies, Alzheimer's and cognitive disorders, schizophrenia, prostate
 CC disease, skeletal or cardiac muscle disorders, pulmonary disorders,
 CC transplant rejection, disorders involving osteoclasts such as
 CC osteoporosis, arthritis or malignancies, digestive/endocrine disorders,
 CC infections and AIDS. The human secreted proteins of the invention are
 CC represented in X37451-X37552 and the encoding nucleic acids are
 CC Sequence 1090 BP; 256 A; 332 C; 295 G; 194 T;

Query Match 2.8%; Score 48; DB 1; Length 1090;
 Best Local Similarity 100.0%; Pred. No. 7.1e-09;
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1666 CTTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1713
 DB 1033 CTTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1080

RESULT 43

N91576
 ID N91576 standard; cDNA; 1239 BP.
 AC N91576;
 DT 28-FEB-1990 (first entry)
 DE Sequence of clone 12R.
 KW cDNA clone; clone 12R; Lol PI; Poaceae; Gramineae; Lolium perenne.
 OS Lolium perenne.
 PN WO8909260-A.

```
PD 05-OCT-1989.
PE 23-MAR-1989: AU0123.
PR 23-MAR-1988: AU-PI7351.
PA (UYME) University of Melbourne.
PI Singh M, Hough T, Theerakulpisut P, Knox RB:
DR WPI: 89-309528/42.
PT Recombinant Lol pi protein - for use in tissue specificity diagnosis and
PT antibody detection.
PS Disclosure: Fig 5.1-5.3: 86pp: English.
CC The cDNA clone 12R contains cDNA encoding Lol PI from Lolium perenne.
SQ Sequence 1239 BP: 290 A; 421 C; 308 G; 220 T;

Query Match
Best Local Similarity 100.0%; Score 47; DB 1; Length 1239;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1667 TTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1713
Db 1185 TTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1231

RESULT 44
N91467
ID N91467 standard: cDNA: 2549 BP.
AC N91467:
DE 05-MAR-1990 (first entry)
DE Sequence of human BP53 from cDNA clone ibp.118.
KW Insulin-like growth factor binding protein; BP53; clone ibp.118.
OS Homo sapiens.
FH Key
FT Location/Qualifiers
FT cds 110..190
FT /*tag= a
FT /*note=Encodes signal sequence."
FT /*tag= b
FT /*note="Encodes mature peptide."

PN W08909268-A.
PD 05-OCT-1989.
PE 10-MAR-1989: U00983.
PR 22-MAR-1988: US-171623.
PA (ROYA-) Royal Prince Alfred Hospital (GETH) Genentech Inc.
PI Baxter KC, Wood WI:
DR WPI: 89-309533/42.
P-PSDB: P92300.
PT DNA encoding insulin-like growth factor circulatory half-life, and as
PT increase insulin-like growth factor circulatory half-life, and as
PT metabolic regulator.
PS Claim 1: page 48: 72pp: English.
CC The larger internal sequence allows hybridisation probes to be obtd. to
CC identify clones with sequences encoding BP53. It is expressed in
CC PKR5-1bp1.1, and used to transform, eg COS cells or human embryonic
CC kidney cells transfected with adenovirus Eladn1b (293s).
SQ Sequence 2549 BP: 671 A; 640 C; 660 G; 578 T;

Query Match
Best Local Similarity 100.0%; Score 47; DB 1; Length 2549;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1667 TTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1713
Db 2474 TTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2520

RESULT 45
Q04690
ID 004690 standard: cDNA: 1834 BP.
AC 004690:
DE 05-OCT-1990 (first entry)
DE Encodes Mammalian amino acid dehydrogenase activating factor-eta
KW Mammalian amino acid dehydrogenase activating protein-eta;
tyrosine dehydrogenase; tryptophan dehydrogenase; ss.
```

```
OS synthetic.
FH key Location/Qualifiers
FT poly_a-signal 1671..1676
FT /*tag= a
FT cds 166..1006
FT /*tag= b

PN J02111796-A.
PD 24-APR-1990.
PE 21-OCT-1988: 264097.
PR 21-OCT-1988: JP-264097.
PA (TOFU) Toa Nenryo Kogyo KK.
DR WPI: 90-169110/22.
P-PSDB: R05084.
PT Amino acid dehydrogenase activating protein -
PT isolated from mammal brain and activates tyrosine dehydrogenase
PT and tryptophan dehydrogenase
PS Disclosure: P: Japanese.
CC Probable error in specification at posn 168. C residue should read
CC G (to give Met codon ATG instead of ATC).
CC Encodes a protein characterised by Mr of 27KD (by SDS-PAGE), iso-
CC electric point of 4.6 and an N-terminal amino acid sequence beginning
CC MelGlyasp-.
CC Protein activates Tyr dehydrogenase and Trp dehydrogenase in the presence
CC of calcium ions and calmodulin-dependent protein phosphatase type II.
CC Useful as a pharmaceutical analysis reagent.
CC See also Q04691 and Q04692.
SQ Sequence 1834 BP: 546 A; 454 C; 471 G; 363 T;

Query Match
Best Local Similarity 100.0%; Score 47; DB 1; Length 1834;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1667 TTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1713
Db 1695 TTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1741

RESULT 46
N80489
ID N80489 standard: DNA: 372 BP.
AC N80489:
DE 16-NOV-1990 (first entry)
DE Cowpea trypsin inhibitor C-terminal coding sequence of PAGCL.
KW Bowman-Birk cowpea trypsin inhibitor (CBTI); insect resistance;
KW PAGCL; ss.
OS Vigna unguiculata.
PN EP-272144-A.
PD 22-JUN-1988.
PE 18-DEC-1987: 311201.
PR 19-DEC-1986: GB-030448.
PR 02-NOV-1987: GB-025610.
PA (AGRI-) Agric Genetics Ltd.
PI Hilder VA, Gatehouse AMR, Boulter D:
DR WPI: 88-169408/25.
PT Recombinant DNA for transfer of specific genes into plants -
PT esp for imparting insect resistance to plants and crops.
PS Disclosure: P: English.
CC PAGCL encodes the C-terminal half of a CPTI. The entire protein
CC is required for insect resistance and so this sequence was used as
CC a probe for in situ colony hybridisation. Full-length coding
CC sequences isolated in this way are used to transfer insect
CC resistance into economically important crops such as maize, rice,
CC soyabeans, vines, sugar beet, cotton, fruit and vegetables.
CC See also P80887-P80889, N80490-N80491 and N82300.
SQ Sequence 372 BP: 148 A; 58 C; 70 G; 96 T;

Query Match
Best Local Similarity 100.0%; Score 47; DB 1; Length 372;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1667 TTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1713
```

```

Db 320 TTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 366
|||||
RESULT 47
Q12023
ID Q12023 standard; cDNA; 7722 BP.
AC Q12023;
DE 13-AUG-1991 (first entry)
DE Norwalk virus genome sense strand
DE gastroenteritis; shellfish; Norwalk virus; detection; ss.
OS Norwalk virus.
FH Key Location/Qualifiers
FT cds 4543..4924
FT /*tag= a
FT /product= RNA-dependent RNA polymerase
PN W09107502.A.
PD 30-MAY-1991.
PF 30-OCT-1990; U06285.
PR 08-NOV-1989; US-433492.
PR 27-APR-1990; US-515993.
PR 27-AUG-1990; US-573509.
PA (PAXU ) BAYLOR COLLEGE MED.
PI Estes WK, Jiang X, Graham DY:
DR WPI: 91-178111/24.
DR P-PSDB: R12298.
PT Double stranded cDNA for Norwalk-type virus detection - comprises
PT fragments large enough to bind Norwalk or related virus genome
PT for identification and characterisation
PS Claim 1: Page 40: 70pp; English.
CC This sequence may contain an additional guanine between G and C at
CC nucleotides 5354 and 5355. The source of cDNA was stool samples from
CC human volunteers administered Norwalk virus. A cDNA library was
CC prepared from nucleic acids extracted from purified viruses.
CC Positive Norwalk virus clone pUCNV-953 was identified from an
CC amplified library (see Q12025): It reacted with post- and pre-
CC infection stool samples from two volunteers. pUCNV-953 was used
CC to identify other clones containing fragments of the Norwalk virus
CC genome. The sequence given here was established in this way.
SO Sequence 7722 BP: 2187 A; 1790 C; 1879 G; 1866 T;

Query Match 2.7%; Score 47; DB 1; Length 7722;
Best Local Similarity 100.0%; Pred. No. 1.1e-08;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1667 TTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1713
|||||
Db 7639 TTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 7685

RESULT 48
Q11760/c
ID Q11760 standard; DNA; 86 BP.
AC Q11760;
DE 22-JUL-1991 (first entry)
DE Self-complementary, T7 promoter hairpin-forming sequence #2.
KW T7 bacteriophage; DNA-dependent RNA polymerase;
KW target sequence amplification; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT stem_loop 1..86
FT /*tag= a
FT /note= "number of T's in loop is 50"
PN EP-427073-A.
PD 15-MAY-1991.
PF 27-OCT-1990; 120650.
PR 09-NOV-1989; US-434372.
PR 23-AUG-1990; US-569991.
PA (MOLE-) MOLECULAR DIAGNOSTI.
PI Dattagupta N:
DR WPI: 91-141658/20.
PT Nucleic acid probe for amplification and detection of target

```

```

PT sequence - capable of forming ligatable hairpin structured
PT promoter and transcribing target sequence, is sensitive and
PT useful in medical diagnosis
PS Claim 2: Page 11: 15pp; English.
CC The sequence is an example of a preferred T7 RNA polymerase
CC promoter for use in the invention. There can be from 2 to 50 T's in
CC the loop region. A probe sequence is ligated to the 3' end of the
CC promoter region. Upon hybridisation of the probe to a target sequence
CC and ligation of the hybridised target sequence to the 5' end of the
CC hairpin-forming sequence, the target sequence can be transcribed by
CC T7 RNA polymerase. This allows the target sequence to be amplified
CC (and detected) using a single oligonucleotide component (c.f. PCR).
CC See also Q11759 and Q11761-Q11764.
SO Sequence 86 BP: 12 A; 6 C; 6 G; 62 T;

Query Match 2.7%; Score 47; DB 1; Length 86;
Best Local Similarity 100.0%; Pred. No. 2.7e-08;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1667 TTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1713
|||||
Db 71 TTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 25

RESULT 49
Q15139
ID Q15139 standard; DNA; 1793 BP.
AC Q15139;
DE 09-MAR-1992 (first entry)
DE Clone pTACC6 encoding the tomato ACC synthase.
KW 1-aminocyclopropane-1-carboxylic acid synthase; fruit ripening; ss.
OS Lycopersicon esculentum c.v. Rutgers.
FH Key Location/Qualifiers
FT cds 91..1548
FT /*tag= a
FT US7579896.A.
PN 12-NOV-1991.
PD 10-SEP-1990; 579896.
PF 10-SEP-1990; US-579896.
PR 10-SEP-1990; US-579896.
PA (USDA ) US SEC OF AGRICULTURE.
PI Theologis A, Sato T:
DR WPI: 91-368895/50.
DR P-PSDB: R15506.
PT DNA encoding ACC synthase - used for control of plant development
PT and for prodn. of ACC synthase, ethylene and ethanol
PS Disclosure: Fig 8: 73pp; English.
CC A lambda gt10 library was constructed using cDNA prepared from polyA
CC RNA isolated from powdered, frozen fruit (i.e. tomatoes).
CC Recombinant phage containing inserts were plated, transferred to
CC a nitrocellulose filter and hybridised to zucchini pACC1 cDNA as
CC probe (see Q15131). A full-length cDNA from tomato, designated
CC pTACC1 was recovered. Additional clones including pTACC6, were
CC isolated using the 3'-end of pTACC1 as a probe.
CC See Q15131-Q15140.
SO Sequence 1793 BP: 631 A; 275 C; 354 G; 533 T;

Query Match 2.7%; Score 47; DB 1; Length 1793;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1667 TTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1713
|||||
Db 1734 TTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1780

RESULT 50
Q15022
ID Q15022 standard; DNA; 1410 BP.
AC Q15022;
DE 16-MAR-1992 (first entry)
DE Hyoscyamine 6 beta-hydroxylase gene.

```

KW Scopolamine; 6-beta-hydroxyhyoscyamine; ss.
 OS Hyoscyamus niger; Location/Qualifiers
 FH Key cds 34..1068 /*tag= a
 FT polyA_site 1344..1410 /*tag= b
 FT J03247277-A.
 PN 05-NOV-1991;
 PD 26-FEB-1990; 042639.
 PF 26-FEB-1990; JP-042639.
 PR (MITC) MITSUI PETROCHEM IND KK.
 PA (MAXF) MAX FACTOR CO. .
 PA WPI: 91-366326/50.
 DR P-PSDB: R15347;
 DR tripane alkaloid synthetase and gene - used for prepn. of
 PI 6-beta-hydroxylase hyoscyamine and scopolamine using
 PI multifunctional enzyme
 PS Claim 1: Fig 12: 12pp: Japanese.
 CC The sequence is that of the hyoscyamine 6 beta-hydroxylase (HH)
 CC gene which encodes a multifunctional enzyme and can be used to prepare
 CC both 6 beta-hydroxyhyoscyamine and scopolamine. Use of recombinant DNA
 CC techniques allows prodn. of large amts. of the protein.
 SO Sequence 1410 BP; 483 A; 236 C; 254 G; 437 T;

Query Match 2.7%; Score 47; DB 1; Length 1410;
 Best Local Similarity 100.0%; Pred. No. 1.5e-08;
 Matches 47: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1567 TTTAAA 1713
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 1342 TTTAAA 1388

Search completed: August 8, 1999, 03:21:46
 Job time: 1630 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 8, 1999, 02:17:30 : Search time 284.88 seconds
(without alignments) 12027.137 Million cell updates/sec

Title: US-09-104-063-3
Sequence: 1737
1 GAATTCACGATGCTGCGC.....CCGCACACACCTGGAATTC 1737

Scoring table: OLIGO_NUC

Searched: 2546578 segs, 986266752 residues

Database :

EST: *
1: em_est1: *
2: em_est2: *
3: em_est3: *
4: em_est4: *
5: em_est5: *
6: em_est6: *
7: em_est7: *
8: em_est8: *
9: em_est9: *
10: em_est10: *
11: em_est11: *
12: em_est12: *
13: em_est13: *
14: em_est14: *
15: em_est15: *
16: em_est16: *
17: em_est17: *
18: em_est18: *
19: em_est19: *
20: em_est20: *
21: em_est21: *
22: em_est22: *
23: em_est23: *
24: em_est24: *
25: em_est25: *
26: em_est26: *
27: em_est27: *
28: em_est28: *
29: em_est29: *
30: em_est30: *
31: em_est31: *
32: em_est32: *
33: em_est33: *
34: em_est34: *
35: em_est35: *
36: em_est36: *
37: em_est37: *
38: em_est38: *
39: em_est39: *
40: em_est40: *
41: em_est41: *
42: em_est42: *
43: em_est43: *
44: em_est44: *
45: em_est45: *
46: em_est46: *
47: em_est47: *
48: em_est48: *
49: em_est49: *
50: em_est50: *
51: em_est51: *
52: em_est52: *
53: em_est53: *

54: em_est32: *
55: em_est33: *
56: em_est24: *
57: em_est25: *
58: em_est26: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB	ID	Description
1	589	33.9	589	40	AA935648	AA935648 op27c09.s
2	561	32.3	599	33	AA411265	AA411265 zv24h06.r
3	539	31.0	539	36	AA610463	AA610463 np97b09.s
4	518	29.8	593	35	AA576017	AA576017 nm57d04.s
5	501	28.8	501	33	AA426644	AA426644 zv47h11.s
6	498	28.7	657	47	AI493618	AI493618 ln39e07.x
7	486	28.0	693	36	AA634211	AA634211 ac72b06.s
8	480	27.6	651	36	AA634201	AA634201 ac72a07.s
9	470	27.1	470	33	AA426566	AA426566 zv47h11.r
10	469	27.0	536	38	AA804282	AA804282 nw30h06.s
11	462	26.6	792	38	AA781110	AA781110 aj23e10.s
12	461	26.5	556	38	AA743645	AA743645 ny24g09.s
13	441	25.4	455	44	AI300807	AI300807 gn47b03.x
14	439	25.3	463	45	AI344724	AI344724 qp04h09.x
15	437	25.2	482	41	AI002547	AI002547 og90e01.s
16	435	25.0	524	44	AI249788	AI249788 gx51c10.x
17	428	24.6	428	30	AA259199	AA259199 nc17b08.r
18	427	24.6	473	44	AI300764	AI300764 gn50c09.x
19	423	24.4	535	41	AI022030	AI022030 ow72a11.x
20	421	24.2	583	29	AA148292	AA148292 zo45e12.s
21	419	24.1	419	44	AI284206	AI284206 q126h07.x
22	409	23.5	409	34	AA490436	AA490436 aa51e04.r
23	409	23.5	465	45	AI370816	AI370816 ta58c10.x
24	403	23.2	509	38	AA760743	AA760743 n210b09.s
25	403	23.2	417	45	AI358446	AI358446 gx19c11.x
26	401	23.1	401	49	AI660355	AI660355 w62b07.x
27	393	22.6	393	33	AA386000	AA386000 EST99758
28	386	22.2	521	34	AA479467	AA479467 zv17f10.r
29	385	22.2	504	43	AI159856	AI159856 qb51c09.x
30	385	22.2	475	44	AI193140	AI193140 q640f12.s
31	375	21.6	531	46	AI434652	AI434652 t134d02.x
32	366	21.1	537	34	AA479357	AA479357 zv17f10.s
33	354	20.4	405	50	AI682902	AI682902 w669d07.x
34	354	20.4	786	37	AA707668	AA707668 z129f12.s
35	347	20.0	786	37	AA284857	AA284857 z122e05.s
36	342	19.7	454	31	AA284857	AA284857 z122e05.s
37	328	18.9	484	31	AA284857	AA284857 z122e05.s
38	308	17.7	308	34	AA501992	AA501992 ne51h08.s
39	299	17.2	361	48	AI610565	AI610565 tp41d08.x
40	299	17.2	407	21	T62636	T62636 yc03f11.s1
41	296	17.0	342	36	AA621854	AA621854 ng19g01.s
42	295	17.0	488	33	AA386001	AA386001 EST98759
43	295	16.9	456	32	AA354393	AA354393 EST62864
44	292	16.8	555	38	AA747545	AA747545 nx85d08.s
45	291	16.8	524	48	AI583169	AI583169 tr98d12.x
46	279	16.1	327	22	R78620	R78620 y174a05.s1
47	277	15.9	277	48	AI582682	AI582682 tn16d04.x
48	275	15.8	427	40	AA934643	AA934643 oo71c08.s
49	274	15.8	347	22	R48597	R48597 vj55e07.s1
50	274	15.8	274	48	AI581657	AI581657 as03h04.x
51	271	15.6	273	32	AA339027	AA339027 EST44092
52	270	15.5	318	32	AA378777	AA378777 EST91512
53	264	15.2	373	36	AA661523	AA661523 ns27f07.s
54	263	15.1	328	50	AI695339	AI695339 wa23d09.x
55	259	14.9	259	37	AA732228	AA732228 n295d12.s
56	255	14.7	411	32	AA354912	AA354912 EST65512
57	248	14.3	337	32	AA357257	AA357257 EST65953
58	238	13.7	344	32	AA354044	AA354044 EST62310

59	228	13.1	382	22	R78657	R78657 y174a05.r1	132	49	2.8	190	40	C93155
60	228	13.1	552	37	AA731086	AA731086 nw68a05.s	133	49	2.8	239	42	AI116374
61	226	13.0	380	32	AA354040	AA354040 EST62304	134	49	2.8	494	44	AI121083
62	224	12.9	533	21	T632491	T632491 vc03f1.r1	135	49	2.8	525	43	AI173457
63	219	12.6	331	32	AA353948	AA353948 EST62146	136	49	2.8	131	43	AI173555
64	218	12.6	270	32	AA34267	AA34267 EST62592	137	49	2.8	312	43	AI182513
65	219	12.6	370	30	AA225739	AA225739 nc17b08.s	138	49	2.8	497	43	AI029685
66	205	11.8	370	20	AA225739	AA225739 nc17b08.s	139	49	2.8	275	44	AI169625
67	201	11.5	307	20	T29035	T29035 EST65954.Hu	140	49	2.8	443	46	AI1397869
68	199	11.5	435	22	R48699	R48699 yj65e02.r1	141	49	2.8	196	48	AI568114
69	192	11.1	398	30	AA248174	AA248174 csg1092.s	142	49	2.8	311	51	AI070779
70	175	10.1	350	23	H29001	H29001 ym31f07.s1	143	48	2.8	433	22	R42275
71	174	10.0	492	29	AA129609	AA129609 z111g10.r1	144	48	2.8	141	24	D19176
72	169	9.7	335	35	AA573161	AA573161 mm51f02.s	145	48	2.8	336	25	N67359
73	166	9.6	200	33	AA436258	AA436258 zv24b06.s	146	48	2.8	397	26	W36406
74	152	8.8	357	47	AI468830	AI468830 t142b08.x	147	48	2.8	327	27	AA008853
75	150	8.6	151	41	AI050023	AI050023 an22a06.x	148	48	2.8	379	27	AA023416
76	128	7.4	437	50	AI675038	AI675038 w222f10.x	149	48	2.8	311	27	W84180
77	122	7.0	157	32	AA357215	AA357215 EST65925	150	48	2.8	357	28	AA098354
78	114	6.6	145	22	R23114	R23114 yh27b12.r1	151	48	2.8	335	28	AA099424
79	112	6.4	214	32	AA353051	AA353051 EST61123	152	48	2.8	464	28	AA099909
80	106	6.1	261	32	AA356857	AA356857 EST65509	153	48	2.8	218	28	AA124210
81	100	6.1	268	32	AA357216	AA357216 EST65926	154	48	2.8	221	28	AA124216
82	100	5.8	100	36	AA634862	AA634862 ab29b06.r1	155	48	2.8	175	29	AA146535
83	99	5.7	249	23	H29103	H29103 ym31f07.r1	156	48	2.8	192	29	AA148291
84	99	5.7	233	31	AA299445	AA299445 EST11954	157	48	2.8	469	29	AA178252
85	96	5.5	411	50	AI683332	AI683332 tw50f10.x	158	48	2.8	564	29	AA192336
86	91	4.7	214	22	R23115	R23115 yh27b12.s1	159	48	2.8	656	30	AA195940
87	75	4.3	254	28	AA091989	AA091989 110224.se	160	48	2.8	375	30	AA216813
88	409	3.1	409	24	N42987	N42987 yy08b05.r1	161	48	2.8	394	30	AA219283
89	53	3.1	385	30	AA221639	AA221639 mw50f03.r1	162	48	2.8	165	30	AA219867
90	53	3.1	136	33	C84892	C84892 C84892 Dict	163	48	2.8	309	30	AA220766
91	53	3.1	331	40	C90379	C90379 C90379 Dict	164	48	2.8	240	30	AA222610
92	53	3.1	250	43	AU053302	AU053302 AU053302	165	48	2.8	335	30	AA238825
93	53	3.1	418	53	HS000060	AI036368 Homo. Sap1	166	48	2.8	308	30	AA240821
94	52	3.0	454	29	AA133892	AA133892 zt85b07.s	167	48	2.8	352	30	AA253985
95	52	3.0	337	30	AA239795	AA239795 mx80a12.r1	168	48	2.8	278	30	AA259893
96	51	2.9	344	30	AA259726	AA259726 v47b11.r1	169	48	2.8	546	30	AA266979
97	51	2.9	400	34	AA470523	AA470523 nc71g11.s	170	48	2.8	288	30	AA267611
98	51	2.9	521	36	AA648402	AA648402 ns22a04.s	171	48	2.8	295	30	AA268010
99	51	2.9	473	37	AA669094	AA669094 aa81e11.s	172	48	2.8	357	30	AA268643
100	51	2.9	539	38	AA743207	AA743207 ny21a08.s	173	48	2.8	240	30	AA268894
101	51	2.9	261	38	AA766932	AA766932 oc88a02.s	174	48	2.8	315	30	AA270150
102	51	2.9	248	38	AA790049	AA790049 vt78g05.r1	175	48	2.8	376	30	AA270475
103	51	2.9	235	47	AI473808	AI473808 tm03e12.x	176	48	2.8	213	30	AA272713
104	51	2.9	344	47	AI529465	AI529465 v47b11.y	177	48	2.8	241	30	AA272932
105	51	2.9	419	49	AI624950	AI624950 ts48c11.x	178	48	2.8	175	31	AA276464
106	51	2.9	132	49	AI642594	AI642594 vv57e01.x	179	48	2.8	447	31	AA285502
107	50	2.9	563	29	AA131231	AA131231 z040h01.r1	180	48	2.8	421	31	AA285554
108	50	2.9	66	30	AA215584	AA215584 zt96c06.s	181	48	2.8	338	31	AA286632
109	50	2.9	265	30	AA244683	AA244683 mx03f02.r1	182	48	2.8	221	31	AA290362
110	50	2.9	270	30	AA250484	AA250484 mx05f08.r1	183	48	2.8	295	33	AA296382
111	50	2.9	227	34	AA488429	AA488429 ab39e09.s	184	48	2.8	309	33	AA433117
112	50	2.9	843	35	AA585438	AA585438 TH842.HTC	185	48	2.8	503	33	AA434832
113	50	2.9	507	38	AA743134	AA743134 ny20g10.s	186	48	2.8	578	34	AA521550
114	50	2.9	93	42	AI096771	AI096771 g806h07.x	187	48	2.8	285	34	AA522107
115	50	2.9	475	42	AI154860	AI154860 ud80a11.r1	188	48	2.8	100	34	AA522914
116	50	2.9	719	49	AF062726	AF062726 AF062726	189	48	2.8	305	35	AA538149
117	50	2.9	627	53	HS000745	AI042559 Homo. Sap1	190	48	2.8	395	35	AA555111
118	49	2.8	134	28	AA118047	AA118047 mm12d11.r1	191	48	2.8	126	35	AA571476
119	49	2.8	358	30	AA222538	AA222538 mw20f07.r1	192	48	2.8	343	35	AA592646
120	49	2.8	229	30	AA254025	AA254025 v110g02.r1	193	48	2.8	309	35	AA592692
121	49	2.8	384	30	AA269935	AA269935 v464d09.r1	194	48	2.8	252	35	AA592892
122	49	2.8	703	35	AA538475	AA538475 vj53e08.r1	195	48	2.8	282	35	AA593292
123	49	2.8	780	35	AA585453	AA585453 TH508.HTC	196	48	2.8	416	35	AA597160
124	49	2.8	302	36	AA607600	AA607600 v048e06.r1	197	48	2.8	240	36	AA607417
125	49	2.8	474	36	AA611263	AA611263 v051b10.r1	198	48	2.8	355	36	AA611773
126	49	2.8	280	37	AA676008	AA676008 TENFO522	199	48	2.8	287	36	AA611938
127	49	2.8	202	37	AA721469	AA721469 v274e07.s	200	48	2.8	181	36	AA616258
128	49	2.8	309	37	AA734289	AA734289 v226e04.r1	201	48	2.8	428	36	AA616585
129	49	2.8	854	38	AA745985	AA745985 ob18h03.s	202	48	2.8	221	36	AA617580
130	49	2.8	246	38	AA789783	AA789783 vt79c09.r1	203	48	2.8	236	36	AA652574
131	49	2.8	220	39	AA822122	AA822122 vp35a12.r1	204	48	2.8	356	36	AA658033

C 205	48	2.8	400	37	AA65612	AA65612 ag55a06.s	C 278	48	2.8	173	46	A1432532	A132552	th30c11.x
C 206	48	2.8	362	37	AA673896	AA673896 vo86g04.i	C 279	48	2.8	372	46	A1432975	A132975	th42a06.x
C 207	48	2.8	405	37	AA680531	AA680531 t3678 Bio	C 280	48	2.8	447	46	A1433657	A133657	tl88a02.x
C 208	48	2.8	497	37	AA682049	AA682049 vu13b08.s	C 281	48	2.8	480	46	A1443503	A143503	sa32f04.x
C 209	48	2.8	139	37	AA689755	AA689755 vs03h03.i	C 282	48	2.8	167	46	A1445620	A1445620	tl08b05.x
C 210	48	2.8	462	37	AA715811	AA715811 nw25b10.s	C 283	48	2.8	351	46	A1446457	A1446457	tl19e05.x
C 211	48	2.8	225	37	AA734226	AA734226 vt26a09.i	C 284	48	2.8	200	47	A1467537	A1467537	ve35f05.x
C 212	48	2.8	533	37	N96086	N96086 12401 CD4.i	C 285	48	2.8	157	47	A1472566	A1472566	tl178b12.x
C 213	48	2.8	213	38	AA202239	AA202239 LD02a23.5	C 286	48	2.8	397	47	A1473347	A173347	fb58g08.x
C 214	48	2.8	357	38	AA744531	AA744531 ny79a06.s	C 287	48	2.8	285	47	A1477670	A1477670	tl16g07.x
C 215	48	2.8	531	38	AA744713	AA744713 nv57b04.s	C 288	48	2.8	293	47	A1480118	A180118	tm33g03.x
C 216	48	2.8	197	38	AA745607	AA745607 nx18c10.s	C 289	48	2.8	349	47	A1493858	A1493858	q251g05.x
C 217	48	2.8	282	38	AA748519	AA748519 ny05f06.s	C 290	48	2.8	368	47	A1498582	A1498582	tn02e02.x
C 218	48	2.8	413	38	AA760851	AA760851 nz14f09.s	C 291	48	2.8	413	47	A1500558	A1500558	tn93g07.x
C 219	48	2.8	354	38	AA761919	AA761919 n242h05.s	C 292	48	2.8	275	47	A1503148	A1503148	vm99e10.x
C 220	48	2.8	518	38	AA766249	AA766249 oa13c08.s	C 293	48	2.8	396	47	A1506810	A1506810	vm61e04.x
C 221	48	2.8	389	38	AA766259	AA766259 oa29a06.s	C 294	48	2.8	578	47	A1535918	A1535918	v1cpr02.A
C 222	48	2.8	258	38	AA778497	AA778497 af20h03.s	C 295	48	2.8	526	47	A1537187	A1537187	tp01d01.x
C 223	48	2.8	202	38	AA792161	AA792161 vt71b11.i	C 296	48	2.8	118	47	A1537516	A1537516	tp04e05.x
C 224	48	2.8	513	38	AA807510	AA807510 ob92a05.s	C 297	48	2.8	169	47	A1538342	A1538342	tp56d11.x
C 225	48	2.8	599	38	AA807691	AA807691 nw31b08.s	C 298	48	2.8	193	47	A1538733	A1538733	tp73f12.x
C 226	48	2.8	441	39	AA814485	AA814485 of42a11.s	C 299	48	2.8	222	47	A1539028	A1539028	tp71d05.x
C 227	48	2.8	168	39	AA816183	AA816183 vp44h07.i	C 300	48	2.8	291	47	A1539219	A1539219	tp55h02.x
C 228	48	2.8	140	39	AA820047	AA820047 ak01f06 M	C 301	48	2.8	188	47	A1540458	A1540458	tg24c01.x
C 229	48	2.8	334	39	AA828415	AA828415 oc46a01.s	C 302	48	2.8	321	48	A1545505	A1545505	fb66d03.x
C 230	48	2.8	348	39	AA830282	AA830282 oc49e02.s	C 303	48	2.8	386	48	A1554818	A1554818	fb73f12.x
C 231	48	2.8	377	39	AA831447	AA831447 oc74f10.s	C 304	48	2.8	279	48	A1558273	A1558273	fb78a09.x
C 232	48	2.8	354	39	AA833819	AA833819 od62b10.s	C 305	48	2.8	306	48	A1558487	A1558487	fb79g10.x
C 233	48	2.8	245	39	AA835970	AA835970 oc80g03.s	C 306	48	2.8	372	48	A1560010	A1560010	fq38b10.x
C 234	48	2.8	341	39	AA837211	AA837211 od24f08.s	C 307	48	2.8	486	48	A1560184	A1560184	tg59h02.x
C 235	48	2.8	217	39	AA847452	AA847452 of07e01.s	C 308	48	2.8	208	48	A1560806	A1560806	tg40b02.x
C 236	48	2.8	531	39	AA870364	AA870364 vq44e05.i	C 309	48	2.8	222	48	A1564259	A1564259	tg75d10.x
C 237	48	2.8	523	39	AA872185	AA872185 ob78a10.s	C 310	48	2.8	383	48	A1566042	A1566042	tn52f07.x
C 238	48	2.8	136	39	C83908	C83908 C83908 Dict	C 311	48	2.8	693	48	A1567501	A1567501	lr89a10.x
C 239	48	2.8	481	39	C84058	C84058 C84058 Dict	C 312	48	2.8	320	48	A1567814	A1567814	lq87a01.x
C 240	48	2.8	212	40	AA904121	AA904121 cq20b11.s	C 313	48	2.8	273	48	A1568338	A1568338	tn68h10.x
C 241	48	2.8	430	40	AA959983	AA959983 vw53d06.s	C 314	48	2.8	337	48	A1577018	A1577018	ui-R-ABD-
C 242	48	2.8	389	40	AA979970	AA979970 MEST4-r3.	C 315	48	2.8	351	48	A1584066	A1584066	ts13c06.x
C 243	48	2.8	262	40	C91078	C91078 C91078 Dict	C 316	48	2.8	141	48	A1584118	A1584118	ts13h06.x
C 244	48	2.8	171	40	C91526	C91526 C91526 Dict	C 317	48	2.8	294	48	A1584353	A1584353	fb92e08.x
C 245	48	2.8	297	41	A1049575	A1049575 an34b09.x	C 318	48	2.8	146	48	A1590535	A1590535	tw11b06.x
C 246	48	2.8	426	41	A1049923	A1049923 an34e08.x	C 319	48	2.8	269	48	A1591566	A1591566	vt25h10.x
C 247	48	2.8	231	42	A1139104	A1139104 qc27g03.x	C 320	48	2.8	250	48	A1601818	A1601818	fc13c08.x
C 248	48	2.8	111	43	A1224373	A1224373 gx05d05.x	C 321	48	2.8	313	48	A1609593	A1609593	tw28d05.x
C 249	48	2.8	344	43	AU029297	AU029297 AU029297	C 322	48	2.8	402	48	A1610822	A1610822	lp38h08.x
C 250	48	2.8	459	43	AU029802	AU029802 AU029802	C 323	48	2.8	212	48	A1612723	A1612723	lp11a10.x
C 251	48	2.8	446	43	AU031165	AU031165 AU031165	C 324	48	2.8	109	49	A1625421	A1625421	ts68e02.x
C 252	48	2.8	421	43	C96980	C96980 C96980 Rice	C 325	48	2.8	584	49	A1625475	A1625475	ly55a06.x
C 253	48	2.8	379	43	C97078	C97078 C97078 Rice	C 326	48	2.8	76	49	A1628214	A1628214	lv22h05.x
C 254	48	2.8	243	43	C97221	C97221 C97221 Rice	C 327	48	2.8	248	49	A1635528	A1635528	ts95c04.x
C 255	48	2.8	301	43	C97413	C97413 C97413 Rice	C 328	48	2.8	354	49	A1636187	A1636187	tz06e07.x
C 256	48	2.8	515	43	C99321	C99321 C99321 Rice	C 329	48	2.8	435	49	A1636309	A1636309	tz23f07.x
C 257	48	2.8	448	43	C99324	C99324 C99324 Rice	C 330	48	2.8	189	49	A1641270	A1641270	fc21g09.x
C 258	48	2.8	338	44	A1256335	A1256335 u184h07.x	C 331	48	2.8	297	49	A1645693	A1645693	ly04c01.x
C 259	48	2.8	523	44	A1262104	A1262104 q228d10.y	C 332	48	2.8	408	49	A1653979	A1653979	qy00c08.A
C 260	48	2.8	350	44	A1280517	A1280517 qu03e04.x	C 333	48	2.8	308	49	A1654015	A1654015	ly04f12.x
C 261	48	2.8	366	44	A1289134	A1289134 gn25b06.x	C 334	48	2.8	186	49	A1655735	A1655735	tl14h03.x
C 262	48	2.8	215	44	A1289400	A1289400 gw32e06.x	C 335	48	2.8	327	49	A1658794	A1658794	tu03a04.x
C 263	48	2.8	374	44	A1316658	A1316658 u116g11.y	C 336	48	2.8	506	49	A1660432	A1660432	we67a11.x
C 264	48	2.8	115	44	AU037787	AU037787 AU037787	C 337	48	2.8	189	49	AU053069	AU053069	tu05j069
C 265	48	2.8	450	45	A1334621	A1334621 lb20e02.x	C 338	48	2.8	290	49	AU057778	AU057778	AU057778
C 266	48	2.8	206	45	A1335214	A1335214 ta93b05.x	C 339	48	2.8	266	49	AU057977	AU057977	AU057977
C 267	48	2.8	450	45	A1342124	A1342124 qt26b12.x	C 340	48	2.8	323	50	A1666935	A1666935	fc43a09.x
C 268	48	2.8	317	45	A1348964	A1348964 ta97h04.x	C 341	48	2.8	403	50	A1669005	A1669005	ty30f03.x
C 269	48	2.8	390	45	A1351737	A1351737 qt03e04.x	C 342	48	2.8	376	50	A1669864	A1669864	wb89g11.x
C 270	48	2.8	333	45	A1360377	A1360377 qy85f01.x	C 343	48	2.8	370	50	A1671931	A1671931	wc26g07.x
C 271	48	2.8	361	45	A1362495	A1362495 qv16h10.x	C 344	48	2.8	224	50	A1672764	A1672764	we58e09.x
C 272	48	2.8	483	45	A1363744	A1363744 qv78e01.x	C 345	48	2.8	340	50	A1674838	A1674838	wc77b07.x
C 273	48	2.8	182	45	A1368579	A1368579 qv61b10.x	C 346	48	2.8	712	50	A1676881	A1676881	tu58f04.x
C 274	48	2.8	277	45	A1371984	A1371984 ta35b10.x	C 347	48	2.8	391	50	A1679209	A1679209	tu22f10.x
C 275	48	2.8	379	45	A1372009	A1372009 ta35e06.x	C 348	48	2.8	220	50	A1679438	A1679438	tu63g11.x
C 276	48	2.8	533	46	A1418970	A1418970 tf38c08.x	C 349	48	2.8	315	50	A1679529	A1679529	tu64b10.x
C 277	48	2.8	351	46	A1427074	A1427074 mn30b04.x	C 350	48	2.8	216	50	A1680369	A1680369	tw62f05.x

C 351	48	2.8	405	50	AI683250	C 424	47	2.7	294	28	AA059479	AA059479	zfe6b04.s
C 352	48	2.8	283	50	AI683492	C 425	47	2.7	170	28	AA070777	AA070777	zm54a08.s
C 353	48	2.8	73	50	AI689096	C 426	47	2.7	140	28	AA087985	AA087985	mo39e12.r
C 354	48	2.8	235	50	AI690663	C 427	47	2.7	302	28	AA102339	AA102339	z191c01.s
C 355	48	2.8	249	50	AI692568	C 428	47	2.7	258	28	AA104461	AA104461	mo55b01.r
C 356	48	2.8	356	50	AI692941	C 429	47	2.7	162	28	AA114771	AA114771	ml14e09.r
C 357	48	2.8	366	50	AI694156	C 430	47	2.7	119	28	AA116956	AA116956	ml29b03.r
C 358	48	2.8	144	50	AI696583	C 431	47	2.7	244	28	AA117061	AA117061	ml29d08.r
C 359	48	2.8	236	50	AI699255	C 432	47	2.7	142	28	AA117240	AA117240	ml19f12.r
C 360	48	2.8	384	50	AT701890	C 433	47	2.7	164	28	AA117464	AA117464	ml22d09.r
C 361	48	2.8	635	50	AT702071	C 434	47	2.7	432	28	AA117775	AA117775	mp19f02.r
C 362	48	2.8	119	50	AT7060519	C 435	47	2.7	134	28	AA119775	AA119775	ml13a03.r
C 363	48	2.8	293	50	F34241	C 436	47	2.7	358	28	AA120562	AA120562	ml13c03.r
C 364	48	2.8	71	50	F37855	C 437	47	2.7	102	28	AA122192	AA122192	ml28e03.s
C 365	48	2.8	395	51	AI711980	C 438	47	2.7	202	28	AA133210	AA133210	ml25902.r
C 366	48	2.8	362	51	AI721560	C 439	47	2.7	102	28	C22148	C22148	ml1ya
C 367	48	2.8	654	51	AI728331	C 440	47	2.7	259	29	AA137934	AA137934	ml01a08.r
C 368	48	2.8	626	51	AI730900	C 441	47	2.7	388	29	AA143304	AA143304	zo37a02.s
C 369	48	2.8	564	51	AI734555	C 442	47	2.7	226	29	AA143316	AA143316	zo37c02.s
C 370	48	2.8	952	53	HSM000564	C 443	47	2.7	314	29	AA144045	AA144045	me65a08.r
C 371	48	2.8	1023	53	HSM001913	C 444	47	2.7	296	29	AA172704	AA172704	me98b10.r
C 372	48	2.8	1135	53	HSM001933	C 445	47	2.7	368	29	AA172938	AA172938	me77c02.r
C 373	48	2.8	273	53	HSM002678	C 446	47	2.7	384	29	AA172938	AA172938	me78c04.r
C 374	48	2.8	172	53	HSM003081	C 447	47	2.7	491	29	AA179186	AA179186	me84e08.r
C 375	48	2.8	147	53	HSM003359	C 448	47	2.7	793	29	AA182270	AA182270	me88c02.r
C 376	48	2.8	465	53	HSM003763	C 449	47	2.7	286	29	AA182270	AA182270	me83f01.r
C 377	48	2.8	190	53	HSM005078	C 450	47	2.7	213	29	AA183287	AA183287	me76h12.r
C 378	48	2.8	790	53	HSM007703	C 451	47	2.7	247	29	AA183305	AA183305	me77c02.r
C 379	48	2.8	650	54	HSM010334	C 452	47	2.7	313	29	AA183047	AA183047	me78c04.r
C 380	48	2.8	471	54	HSM010561	C 453	47	2.7	212	29	AA184945	AA184945	me84e08.r
C 381	48	2.8	86	54	HSM010597	C 454	47	2.7	309	29	AA184997	AA184997	me88c02.r
C 382	48	2.8	249	54	HSM011323	C 455	47	2.7	354	29	AA185015	AA185015	me88d09.r
C 383	48	2.8	264	54	HSM011474	C 456	47	2.7	239	29	AA185060	AA185060	me89g11.r
C 384	48	2.8	454	54	HSM011933	C 457	47	2.7	304	29	AA185352	AA185352	me91a02.r
C 385	48	2.8	285	54	HSM012002	C 458	47	2.7	489	29	AA185434	AA185434	me51d02.r
C 386	47	2.7	151	20	T04695	C 459	47	2.7	370	29	AA186025	AA186025	me35e04.r
C 387	47	2.7	135	20	T25552	C 460	47	2.7	381	29	AA186144	AA186144	me71907.r
C 388	47	2.7	175	20	T28421	C 461	47	2.7	171	29	AA186960	AA186960	me90a10.r
C 389	47	2.7	413	20	T49724	C 462	47	2.7	155	29	AA193813	AA193813	rs08g10.r
C 390	47	2.7	75	21	R10067	C 463	47	2.7	577	29	AA194303	AA194303	z004e11.s
C 391	47	2.7	109	21	T69241	C 464	47	2.7	397	29	U78999	U78999	me8999.Me
C 392	47	2.7	247	22	H12358	C 465	47	2.7	746	30	AA195998	AA195998	zp99f05.s
C 393	47	2.7	315	22	R36363	C 466	47	2.7	500	30	AA196541	AA196541	zq08e10.s
C 394	47	2.7	228	22	R43224	C 467	47	2.7	342	30	AA198180	AA198180	mv49909.r
C 395	47	2.7	359	23	R80916	C 468	47	2.7	188	30	AA198259	AA198259	mv47f10.r
C 396	47	2.7	359	23	H42657	C 469	47	2.7	303	30	AA198866	AA198866	mv42c07.r
C 397	47	2.7	159	24	H79905	C 470	47	2.7	419	30	AA199403	AA199403	mv39h11.r
C 398	47	2.7	124	24	N25033	C 471	47	2.7	282	30	AA210320	AA210320	mv72a06.r
C 399	47	2.7	61	24	N33110	C 472	47	2.7	286	30	AA210335	AA210335	mv72h03.r
C 400	47	2.7	59	24	N39043	C 473	47	2.7	194	30	AA210586	AA210586	mv71b03.r
C 401	47	2.7	299	24	N41289	C 474	47	2.7	119	30	AA212045	AA212045	mo90g08.r
C 402	47	2.7	350	25	N43396	C 475	47	2.7	100	30	AA212139	AA212139	mo90d11.r
C 403	47	2.7	294	25	N63391	C 476	47	2.7	505	30	AA214091	AA214091	zr58h02.s
C 404	47	2.7	200	25	N97837	C 477	47	2.7	230	30	AA215267	AA215267	mu97c05.r
C 405	47	2.7	128	25	N98038	C 478	47	2.7	258	30	AA216888	AA216888	mu93c06.r
C 406	47	2.7	276	25	N98166	C 479	47	2.7	351	30	AA216946	AA216946	mv75g10.r
C 407	47	2.7	352	25	W14029	C 480	47	2.7	333	30	AA217935	AA217935	mv56e10.r
C 408	47	2.7	306	26	W18492	C 481	47	2.7	381	30	AA220788	AA220788	mv69f04.r
C 409	47	2.7	355	26	W33314	C 482	47	2.7	379	30	AA222625	AA222625	mv77g03.r
C 410	47	2.7	194	26	W38553	C 483	47	2.7	363	30	AA230383	AA230383	mv73a12.r
C 411	47	2.7	450	26	W52670	C 484	47	2.7	329	30	AA230802	AA230802	mu13e02.r
C 412	47	2.7	285	26	W69041	C 485	47	2.7	327	30	AA237194	AA237194	mv96f08.r
C 413	47	2.7	468	26	W73128	C 486	47	2.7	360	30	AA237826	AA237826	mx29d06.r
C 414	47	2.7	92	26	W74529	C 487	47	2.7	351	30	AA238147	AA238147	mx28h02.r
C 415	47	2.7	329	26	W84077	C 488	47	2.7	240	30	AA241764	AA241764	mw26g09.r
C 416	47	2.7	345	26	W91679	C 489	47	2.7	126	30	AA241897	AA241897	mw26g09.r
C 417	47	2.7	122	27	AA033323	C 490	47	2.7	452	30	AA242021	AA242021	mx22e02.r
C 418	47	2.7	360	27	AA050541	C 491	47	2.7	326	30	AA242229	AA242229	mw26f02.r
C 419	47	2.7	535	27	AA056832	C 492	47	2.7	247	30	AA242422	AA242422	mx26a11.r
C 420	47	2.7	240	27	C08964	C 493	47	2.7	278	30	AA250004	AA250004	mw59a10.r
C 421	47	2.7	368	27	W82092	C 494	47	2.7	240	30	AA250343	AA250343	mx59p11.r
C 422	47	2.7	382	27	W99042	C 495	47	2.7	304	30	AA253695	AA253695	mx85c03.r
C 423	47	2.7	510	28	AA058861	C 496	47	2.7	301	30	AA254104	AA254104	va08f05.r

497	47	182	30	AA254324	AA254324	val15d04.r	C	570	261	34	AA481167	AA481167	aa34h09.s
498	47	217	30	AA254413	AA254413	val16b03.r	C	571	550	34	AA496558	AA496558	vf96g01.s
499	47	293	30	AA254562	AA254562	val13c06.r	C	572	106	34	AA502794	AA502794	ne43g08.s
500	47	297	30	AA259598	AA259598	val1b01.r	C	573	75	34	AA503384	AA503384	ne45b11.s
501	47	178	30	AA259635	AA259635	va47f03.r	C	574	291	34	AA505511	AA505511	nh84e11.s
502	47	347	30	AA259742	AA259742	va47f12.r	C	575	313	34	AA505529	AA505529	nh84h02.s
503	47	127	30	AA260102	AA260102	va52a09.r	C	576	222	34	AA512115	AA512115	vf16b10.r
504	47	195	30	AA260691	AA260691	va01g04.r	C	577	324	34	AA512963	AA512963	
505	47	245	30	AA260795	AA260795	va21g03.r	C	578	355	34	AA512985	AA512985	
506	47	227	30	AA261113	AA261113	m265h09.r	C	579	258	34	AA516371	AA516371	
507	47	222	30	AA261246	AA261246	va51h04.r	C	580	233	34	AA517983	AA517983	
508	47	225	30	AA261593	AA261593	m265a08.r	C	581	88	34	AA518548	AA518548	
509	47	338	30	AA265772	AA265772	m273b03.r	C	582	598	34	AA521299	AA521299	
510	47	368	30	AA266673	AA266673	m267a05.r	C	583	268	34	AA521767	AA521767	
511	47	295	30	AA266609	AA266609	m267f10.r	C	584	410	34	AA522073	AA522073	
512	47	239	30	AA266700	AA266700	m296b08.r	C	585	365	34	AA522097	AA522097	
513	47	247	30	AA266722	AA266722	m298f05.r	C	586	545	34	AA525540	AA525540	
514	47	261	30	AA267108	AA267108	m297d09.r	C	587	408	34	AA525735	AA525735	
515	47	400	30	AA267358	AA267358	m291d09.r	C	588	319	34	AA529372	AA529372	
516	47	295	30	AA267875	AA267875	va18g08.r	C	589	532	34	AA529509	AA529509	
517	47	326	30	AA268028	AA268028	vb08h07.r	C	590	284	34	AA529805	AA529805	
518	47	356	30	AA268084	AA268084	va89h01.r	C	591	232	34	AA529851	AA529851	
519	47	380	30	AA268844	AA268844	va41d11.r	C	592	523	34	C24342	C24342	
520	47	301	30	AA268600	AA268600	va43d08.r	C	593	241	34	C24389	C24389	
521	47	359	30	AA268972	AA268972	va99a06.r	C	594	332	34	C24395	C24395	
522	47	287	30	AA269333	AA269333	va41b08.r	C	595	70	35	AA537101	AA537101	
523	47	213	30	AA269395	AA269395	va42h02.r	C	596	340	35	AA538496	AA538496	
524	47	401	30	AA269763	AA269763	va56f12.r	C	597	352	35	AA538916	AA538916	
525	47	327	30	AA269806	AA269806	va63d03.r	C	598	653	35	AA538951	AA538951	
526	47	196	30	AA269952	AA269952	va64h11.r	C	599	675	35	AA541878	AA541878	
527	47	320	30	AA270173	AA270173	va59a11.r	C	600	704	35	AA541933	AA541933	
528	47	285	30	AA270274	AA270274	va63d08.r	C	601	716	35	AA541939	AA541939	
529	47	386	30	AA270384	AA270384	va63d12.r	C	602	271	35	AA543164	AA543164	
530	47	403	30	AA270895	AA270895	vb72e11.r	C	603	710	35	AA545680	AA545680	
531	47	319	30	AA270951	AA270951	va70d10.r	C	604	552	35	AA552445	AA552445	
532	47	298	30	AA271200	AA271200	vb79d11.r	C	605	249	35	AA555735	AA555735	
533	47	367	30	AA271960	AA271960	va73c02.r	C	606	499	35	AA557132	AA557132	
534	47	267	30	AA271992	AA271992	vb79g05.r	C	607	104	35	AA560851	AA560851	
535	47	415	30	AA272436	AA272436	va80a06.r	C	608	224	35	AA562295	AA562295	
536	47	311	30	AA272760	AA272760	va42a07.r	C	609	243	35	AA571665	AA571665	
537	47	306	30	AA273250	AA273250	vc01a03.r	C	610	111	35	AA574881	AA574881	
538	47	364	30	AA273444	AA273444	vc01d10.r	C	611	274	35	AA575541	AA575541	
539	47	314	30	AA273940	AA273940	vb99f04.r	C	612	354	35	AA575800	AA575800	
540	47	27	30	AA274084	AA274084	vb96f12.r	C	613	451	35	AA575922	AA575922	
541	47	1104	30	AA2775	AA2775	C22775	C	614	87	35	AA579618	AA579618	
542	47	227	31	AA275358	AA275358	vc07d01.r	C	615	328	35	AA583169	AA583169	
543	47	427	31	AA275763	AA275763	vc14f09.r	C	616	102	35	AA587120	AA587120	
544	47	242	31	AA275879	AA275879	vc25g06.r	C	617	185	35	AA589036	AA589036	
545	47	238	31	AA277536	AA277536	vb91b07.r	C	618	222	35	AA589218	AA589218	
546	47	368	31	AA279293	AA279293	zs84a11.s	C	619	146	35	AA590017	AA590017	
547	47	303	31	AA285391	AA285391	va67a10.r	C	620	278	35	AA590062	AA590062	
548	47	271	31	AA285459	AA285459	vb90h10.r	C	621	240	35	AA590219	AA590219	
549	47	210	31	AA285469	AA285469	vc32g04.r	C	622	365	35	AA590672	AA590672	
550	47	319	31	AA285492	AA285492	vb82g09.r	C	623	357	35	AA590688	AA590688	
551	47	263	31	AA285774	AA285774	vb82g08.r	C	624	202	35	AA590931	AA590931	
552	47	416	31	AA288487	AA288487	vb39g10.r	C	625	197	35	AA592391	AA592391	
553	47	197	31	AA290346	AA290346	vb33c09.r	C	626	326	35	AA596390	AA596390	
554	47	155	32	AA355098	AA355098	EST63446.r	C	627	126	35	AA597246	AA597246	
555	47	352	33	AA390106	AA390106	vb28d08.r	C	628	351	35	C24633	C24633	
556	47	357	33	AA390238	AA390238	CPEST.126	C	629	282	35	C24639	C24639	
557	47	307	33	AA396391	AA396391	vb40h11.r	C	630	255	35	C24652	C24652	
558	47	264	33	AA397033	AA397033	mx84f09.r	C	631	288	35	C24675	C24675	
559	47	243	33	AA414924	AA414924	vc50g04.r	C	632	498	35	C25597	C25597	
560	47	590	33	AA415124	AA415124	MG0087.RC	C	633	504	35	C25617	C25617	
561	47	329	33	AA427328	AA427328	ve24b05.r	C	634	412	35	C25619	C25619	
562	47	203	33	AA449768	AA449768	zx07h01.s	C	635	510	35	C25685	C25685	
563	47	284	33	AA450824	AA450824	vf78h05.r	C	636	91	35	C25755	C25755	
564	47	539	33	C23689	C23689	C3668.Dict	C	637	363	35	C29527	C29527	
565	47	193	33	C23733	C23733	C3733.Dict	C	638	332	36	AA602414	AA602414	
566	47	268	34	AA473895	AA473895	vb30g12.r	C	639	395	36	AA604005	AA604005	
567	47	523	34	AA473081	AA473081	vd44b07.r	C	640	305	36	AA605362	AA605362	
568	47	558	34	AA479803	AA479803	zua3a07.r	C	641	321	36	AA605942	AA605942	
569	47	256	34	AA481053	AA481053	aa29d04.s	C	642					

AA481167	aa34h09.s
AA496658	vf96g01.s
AA502794	ne43g08.s
AA503384	ne45b11.s
AA505511	nh84e11.s
AA505529	nh84h02.s
AA512115	vf16b10.r
AA512963	nj16m01.s
AA512985	nj16d01.s
AA516371	nf59f11.s
AA517983	vf118b09.r
AA518548	vf17c09.r
AA521299	aa79f05.s
AA521767	vf12e09.r
AA522073	vf16e05.r
AA522097	vf16g07.r
AA525540	nf165b07.s
AA525735	TENFO042
AA529372	vf139e03.r
AA529509	vf140f02.r
AA529805	vf12e09.r
AA529851	vf11g11.r
AA529959	vf154b09.r
AA541878	vf54h03.r
AA541933	vf58g03.r
AA541939	vf58c11.r
AA543164	vf181g06.r
AA545680	vk01g01.r
AA552445	nk15d11.s
AA555735	vk10c10.r
AA557132	n17a12.s
AA560851	vk94f08.r
AA562295	vk97a12.r
AA571665	vm09e04.r
AA574881	vm33a12.r
AA575541	vm03c10.r
AA575800	vm63a02.r
AA575922	nm56a11.s
AA579618	nm79c06.s
AA583169	nm37d01.s
AA587120	nm70b05.s
AA589036	vi65d01.r
AA589218	vk24f01.r
AA590017	vm22e09.r
AA590062	vm57d05.r
AA590219	vn56f12.r
AA590672	vn58b12.r
AA590688	vn58g11.r
AA590931	vm21h01.r
AA592391	vn61f03.r
AA596390	vo38c07.r
AA597246	vo28d05.r
C24633	C24633.Dict
C24639	C24639.Dict
C24652	C24652.Dict
C24675	C24675.Dict
C25597	C25597.Dict
C25617	C25617.Dict
C25619	C25619.Dict
C25685	C25685.Dict
C25755	C25755.Dict
C29527	HSZ98527.DK
AA602414	no30f05.s
AA605362	vc47g09.r
AA605942	vm39d12.r

643	47	2.7	231	36	AA607037	vm95b06.r	C 716	47	2.7	565	38	AA748729	ny02h03.s
644	47	2.7	59	36	AA607050	vm95d06.r	C 717	47	2.7	428	38	AA749349	AA749349 nx99f03.s
645	47	2.7	264	36	AA607075	vm95h03.r	C 718	47	2.7	591	38	AA752896	AA752896 97AS0589
646	47	2.7	142	36	AA607117	vm92b03.r	C 719	47	2.7	521	38	AA753389	AA753389 97BS0012
647	47	2.7	271	36	AA607299	vm96d09.r	C 720	47	2.7	656	38	AA753511	AA753511 96BS0678
648	47	2.7	234	36	AA607418	vm04a12.r	C 721	47	2.7	490	38	AA753566	AA753566 97BS0038
649	47	2.7	131	36	AA607427	vm04i12.r	C 722	47	2.7	532	38	AA753611	AA753611 97BS0091
650	47	2.7	100	36	AA607932	vm39e04.r	C 723	47	2.7	548	38	AA753634	AA753634 97BS0116
651	47	2.7	258	36	AA608073	vm42e10.r	C 724	47	2.7	608	38	AA753645	AA753645 97BS0127
652	47	2.7	167	36	AA608256	vm61b09.r	C 725	47	2.7	561	38	AA753690	AA753690 97BS0181
653	47	2.7	419	36	AA608276	vm61b09.r	C 726	47	2.7	526	38	AA753741	AA753741 97BS0242
654	47	2.7	127	36	AA608316	vm85e03.r	C 727	47	2.7	591	38	AA753847	AA753847 97BS0279
655	47	2.7	331	36	AA611505	vm08b08.r	C 728	47	2.7	592	38	AA753866	AA753866 97BS0307
656	47	2.7	270	36	AA611723	vm08f08.r	C 729	47	2.7	488	38	AA753896	AA753896 97BS0342
657	47	2.7	377	36	AA611806	vm03f06.r	C 730	47	2.7	578	38	AA753952	AA753952 97BS0408
C 658	47	2.7	498	36	AA614244	np09g04.s	C 731	47	2.7	175	38	AA755151	AA755151 v91a10.r
659	47	2.7	345	36	AA615335	vm65a10.r	C 732	47	2.7	323	38	AA755717	AA755717 vm12h11.r
660	47	2.7	397	36	AA616794	vm68b03.r	C 733	47	2.7	197	38	AA756676	AA756676 vm56g09.r
661	47	2.7	202	36	AA617056	vm12c06.r	C 734	47	2.7	339	38	AA760884	AA760884 n218d01.s
662	47	2.7	171	36	AA617507	v177b02.r	C 735	47	2.7	483	38	AA761545	AA761545 n223a09.s
663	47	2.7	340	36	AA619045	vm67a12.r	C 736	47	2.7	412	38	AA761573	AA761573 n23d08.s
C 664	47	2.7	470	36	AA622416	nm45f02.s	C 737	47	2.7	240	38	AA761608	AA761608 n28b05.s
665	47	2.7	397	36	AA624123	vm75h03.s	C 738	47	2.7	325	38	AA764206	AA764206 vm45e10.r
666	47	2.7	262	36	AA638541	vm054e02.r	C 739	47	2.7	341	38	AA764377	AA764377 vp08c12.r
667	47	2.7	245	36	AA638677	vm056b09.r	C 740	47	2.7	365	38	AA764946	AA764946 n260f01.s
668	47	2.7	278	36	AA638755	vm94h11.r	C 741	47	2.7	140	38	AA765198	AA765198 n262b04.s
C 669	47	2.7	136	36	AA641818	ns19f08.s	C 742	47	2.7	344	38	AA765346	AA765346 oa06d06.s
C 670	47	2.7	557	36	AA643953	np42h06.s	C 743	47	2.7	411	38	AA765751	AA765751 oa07g08.s
C 671	47	2.7	499	36	AA648480	ns33d10.s	C 744	47	2.7	405	38	AA765849	AA765849 oa25f06.s
C 672	47	2.7	447	36	AA651857	ns38f01.s	C 745	47	2.7	224	38	AA767009	AA767009 oa42a03.s
C 673	47	2.7	580	36	AA653370	ag65f07.s	C 746	47	2.7	103	38	AA767924	AA767924 oa60a03.s
C 674	47	2.7	509	36	AA653460	ag66c12.s	C 747	47	2.7	473	38	AA782332	AA782332 ac26e08.s
C 675	47	2.7	387	36	AA653587	ag61g09.s	C 748	47	2.7	284	38	AA786399	AA786399 15d01a1.f
C 676	47	2.7	323	36	AA662155	ns66b09.s	C 749	47	2.7	190	38	AA790030	AA790030 vb80e06.r
C 677	47	2.7	360	36	C62130	C62130 yu11	C 750	47	2.7	366	38	AA793183	AA793183 vp29e06.r
678	47	2.7	479	37	AA673883	vm08f02.r	C 751	47	2.7	268	38	AA794469	AA794469 vm66d04.r
679	47	2.7	308	37	AA674843	vq51a05.r	C 752	47	2.7	148	38	AA794691	AA794691 vm63c05.r
680	47	2.7	304	37	AA674852	vq51b02.r	C 753	47	2.7	180	38	AA795084	AA795084 vm09f11.r
681	47	2.7	258	37	AA674905	vq57c10.r	C 754	47	2.7	266	38	AA795884	AA795884 vt28g06.r
682	47	2.7	201	37	AA689105	vm06f09.r	C 755	47	2.7	434	38	AA796766	AA796766 vp27c11.r
C 683	47	2.7	451	37	AA689584	nm66a03.s	C 756	47	2.7	609	38	AA796779	AA796779 vp31a05.r
684	47	2.7	413	37	AA691245	vm14d03.r	C 757	47	2.7	211	38	AA796787	AA796787 vp31b04.r
685	47	2.7	401	37	AA692516	vt59f12.r	C 758	47	2.7	378	38	AA796807	AA796807 vp31d09.r
686	47	2.7	254	37	AA697665	HL02953.5	C 759	47	2.7	222	38	AA797898	AA797898 vm33g04.r
687	47	2.7	557	37	AA701785	PM1FG_139	C 760	47	2.7	252	38	AA804541	AA804541 ns28d07.s
688	47	2.7	623	37	AA701797	PM1FG_150	C 761	47	2.7	284	38	AA804843	AA804843 oc44e06.s
689	47	2.7	330	37	AA709867	vt35a08.r	C 762	47	2.7	379	38	AA805434	AA805434 oc13h11.s
690	47	2.7	276	37	AA710288	vt53a01.r	C 763	47	2.7	337	38	AA805964	AA805964 oc20c05.s
691	47	2.7	384	37	AA710451	vt42f07.r	C 764	47	2.7	435	38	AA806378	AA806378 oc22g02.s
692	47	2.7	175	37	AA710467	vt42h10.r	C 765	47	2.7	218	38	AA806605	AA806605 ob68g03.s
693	47	2.7	240	37	AA710521	vt50f03.r	C 766	47	2.7	297	38	AA806720	AA806720 ob87e04.s
694	47	2.7	357	37	AA710569	vt43c05.r	C 767	47	2.7	427	38	AA806757	AA806757 ob91c06.s
695	47	2.7	312	37	AA711183	vt56g09.r	C 768	47	2.7	437	38	AA807363	AA807363 of50d04.s
C 696	47	2.7	165	37	AA713781	nm70a12.s	C 769	47	2.7	269	38	AA808112	AA808112 of51a07.s
697	47	2.7	382	37	AA727832	vp33d10.r	C 770	47	2.7	133	38	AA808175	AA808175 of51h12.s
698	47	2.7	367	37	AA727836	vp33e02.r	C 771	47	2.7	308	38	AA810226	AA810226 od14a12.s
699	47	2.7	391	37	AA727881	vp34a11.r	C 772	47	2.7	437	38	AA810677	AA810677 oa73a09.s
C 700	47	2.7	391	37	AA731184	nm56e12.s	C 773	47	2.7	474	38	AA811284	AA811284 ob68f04.s
C 701	47	2.7	436	37	AA731241	nm59f04.s	C 774	47	2.7	361	39	AA814343	AA814343 n206g07.s
C 702	47	2.7	151	37	AA731711	nm70c03.s	C 775	47	2.7	320	39	AA814517	AA814517 of46d12.s
C 703	47	2.7	618	38	AA738219	ny16b01.s	C 776	47	2.7	185	39	AA814691	AA814691 of43e12.s
C 704	47	2.7	463	38	AA738499	ny23a01.s	C 777	47	2.7	206	39	AA815526	AA815526 vp28c02.r
C 705	47	2.7	263	38	AA740250	ny79d07.s	C 778	47	2.7	639	39	AA816049	AA816049 vt14h02.r
C 706	47	2.7	377	38	AA741286	n201d01.s	C 779	47	2.7	508	39	AA822925	AA822925 vp30e11.r
C 707	47	2.7	203	38	AA745012	ny20c12.s	C 780	47	2.7	563	39	AA822944	AA822944 vp30g09.r
C 708	47	2.7	294	38	AA745048	ny16b01.s	C 781	47	2.7	299	39	AA823385	AA823385 vp39a08.r
C 709	47	2.7	347	38	AA745387	ny23a01.s	C 782	47	2.7	335	39	AA830396	AA830396 oc56g11.s
C 710	47	2.7	933	38	AA744557	ny79d07.s	C 783	47	2.7	246	39	AA830421	AA830421 oc51a01.s
C 711	47	2.7	411	38	AA744668	ny26a02.s	C 784	47	2.7	285	39	AA830431	AA830431 oc51b05.s
C 712	47	2.7	330	38	AA745069	np73g12.s	C 785	47	2.7	299	39	AA830709	AA830709 oc55h06.s
C 713	47	2.7	607	38	AA745959	ob18e09.s	C 786	47	2.7	473	39	AA830815	AA830815 oc56e10.s
C 714	47	2.7	395	38	AA748697	oa57h06.s	C 787	47	2.7	332	39	AA831948	AA831948 oc90d02.s
C 715	47	2.7	406	38	AA748698	oa57h07.s	C 788	47	2.7	223	39	AA833453	AA833453 oc92g11.s

789	47	2.7	331	39	AA833305	AA833305	od65e08.r	862	47	2.7	450	40	C89917	C89917	Dict
C 750	47	2.7	440	39	AA834534	AA834534	od65e06.s	863	47	2.7	172	40	C89934	C89934	Dict
C 791	47	2.7	242	35	AA835168	AA835168	od16e07.s	864	47	2.7	241	40	C90027	C90027	Dict
C 792	47	2.7	247	39	AA836471	AA836471	od38b12.s	865	47	2.7	270	40	C90071	C90071	Dict
C 793	47	2.7	355	39	AA837391	AA837391	od41e12.s	866	47	2.7	270	40	C90078	C90078	Dict
C 794	47	2.7	517	39	AA837930	AA837930	oe93a09.s	867	47	2.7	217	40	C90121	C90121	Dict
C 795	47	2.7	298	39	AA838319	AA838319	oe98b04.s	868	47	2.7	301	40	C90156	C90156	Dict
C 796	47	2.7	409	39	AA847917	AA847917	od39f01.s	869	47	2.7	621	40	C90408	C90408	Dict
C 797	47	2.7	781	39	AA849835	AA849835	EST192f02	870	47	2.7	196	40	C90466	C90466	Dict
C 798	47	2.7	483	39	AA857969	AA857969	oe93h06.s	871	47	2.7	449	40	C90501	C90501	Dict
C 799	47	2.7	198	39	AA862606	AA862606	oh07b04.s	872	47	2.7	385	40	C90503	C90503	Dict
C 800	47	2.7	204	39	AA862660	AA862660	oh45c03.s	873	47	2.7	442	40	C90506	C90506	Dict
C 801	47	2.7	329	39	AA871153	AA871153	vg43g07.r	874	47	2.7	216	40	C90718	C90718	Dict
C 802	47	2.7	509	39	AA871931	AA871931	vg43g08.r	875	47	2.7	604	40	C90785	C90785	Dict
C 803	47	2.7	210	39	AA872507	AA872507	oi10e11.s	876	47	2.7	264	40	C90816	C90816	Dict
C 804	47	2.7	220	39	AA881672	AA881672	vx21d09.r	877	47	2.7	374	40	C90828	C90828	Dict
C 805	47	2.7	574	39	AA897798	AA897798	NCP4H517	878	47	2.7	403	40	C90829	C90829	Dict
C 806	47	2.7	631	39	AA897888	AA897888	NCC3C9T7	879	47	2.7	198	40	C90840	C90840	Dict
C 807	47	2.7	409	39	AA897891	AA897891	NCC4A6T7	880	47	2.7	686	40	C90855	C90855	Dict
C 808	47	2.7	586	39	AA897904	AA897904	NCP4G1T7	881	47	2.7	348	40	C90856	C90856	Dict
C 809	47	2.7	510	39	AA897913	AA897913	NCC1F1T7	882	47	2.7	340	40	C90858	C90858	Dict
C 810	47	2.7	480	39	AA897914	AA897914	NCC1F1T7	883	47	2.7	373	40	C90878	C90878	Dict
C 811	47	2.7	574	39	AA897926	AA897926	NCC2E8T7	884	47	2.7	390	40	C90883	C90883	Dict
C 812	47	2.7	547	39	AA898229	AA898229	NCC2B9T7	885	47	2.7	309	40	C90935	C90935	Dict
C 813	47	2.7	557	39	AA898385	AA898385	NCC6D4T7	886	47	2.7	308	40	C90985	C90985	Dict
C 814	47	2.7	546	39	AA898463	AA898463	NCC5H12T7	887	47	2.7	256	40	C90988	C90988	Dict
C 815	47	2.7	493	39	AA898589	AA898589	NCC3D1T7	888	47	2.7	643	40	C90989	C90989	Dict
C 816	47	2.7	586	39	AA898674	AA898674	NCM7G7T7	889	47	2.7	521	40	C91004	C91004	Dict
C 817	47	2.7	513	39	AA901562	AA901562	NCM6A1T7	890	47	2.7	276	40	C91123	C91123	Dict
C 818	47	2.7	491	39	AA901781	AA901781	NCP4E1T7	891	47	2.7	475	40	C91125	C91125	Dict
C 819	47	2.7	523	39	AA901860	AA901860	NCP3F4T7	892	47	2.7	487	40	C91130	C91130	Dict
C 820	47	2.7	624	39	AA901889	AA901889	NCP2E6T7	893	47	2.7	252	40	C91203	C91203	Dict
C 821	47	2.7	647	39	AA901984	AA901984	NCM8G8T7	894	47	2.7	405	40	C91306	C91306	Dict
C 822	47	2.7	264	39	C83835	C83835	Dict	895	47	2.7	267	40	C91337	C91337	Dict
C 823	47	2.7	423	39	C83840	C83840	Dict	896	47	2.7	166	40	C91511	C91511	Dict
C 824	47	2.7	357	39	C83928	C83928	Dict	897	47	2.7	430	40	C91532	C91532	Dict
C 825	47	2.7	211	39	C84164	C84164	Dict	898	47	2.7	194	40	C92022	C92022	Dict
C 826	47	2.7	166	39	C84166	C84166	Dict	899	47	2.7	298	40	C92188	C92188	Dict
C 827	47	2.7	350	39	C84677	C84677	Dict	900	47	2.7	630	40	C92364	C92364	Dict
C 828	47	2.7	331	39	C84704	C84704	Dict	901	47	2.7	607	40	C92385	C92385	Dict
C 829	47	2.7	272	39	C84782	C84782	Dict	902	47	2.7	277	40	C92705	C92705	Dict
C 830	47	2.7	235	39	C84788	C84788	Dict	903	47	2.7	414	40	C92768	C92768	Dict
C 831	47	2.7	578	39	C84827	C84827	Dict	904	47	2.7	362	40	C92779	C92779	Dict
C 832	47	2.7	341	39	C84867	C84867	Dict	905	47	2.7	225	40	C93081	C93081	Dict
C 833	47	2.7	149	40	AA908294	AA908294	oq33f11.s	906	47	2.7	187	40	C93096	C93096	Dict
C 834	47	2.7	504	40	AA908482	AA908482	oq82c12.s	907	47	2.7	216	40	C93106	C93106	Dict
C 835	47	2.7	815	40	AA910956	AA910956	oX85h11.s	908	47	2.7	216	40	C93106	C93106	Dict
C 836	47	2.7	195	40	AA911767	AA911767	oq19c07.s	909	47	2.7	625	40	C93193	C93193	Dict
C 837	47	2.7	469	40	AA915056	AA915056	vz01h11.r	910	47	2.7	343	40	C93249	C93249	Dict
C 838	47	2.7	172	40	AA916133	AA916133	oq32b09.s	911	47	2.7	347	40	C93297	C93297	Dict
C 839	47	2.7	176	40	AA923096	AA923096	oK91c09.s	912	47	2.7	108	41	A1010522	A1010522	Dict
C 840	47	2.7	95	40	AA928539	AA928539	om73g01.s	913	47	2.7	294	41	A1011635	A1011635	Dict
C 841	47	2.7	426	40	AA933083	AA933083	om85f10.s	914	47	2.7	227	41	A1016644	A1016644	Dict
C 842	47	2.7	205	40	AA937566	AA937566	oq79d12.s	915	47	2.7	393	41	A1025513	A1025513	Dict
C 843	47	2.7	361	40	AA937574	AA937574	oq79f06.s	916	47	2.7	455	41	A1039104	A1039104	Dict
C 844	47	2.7	465	40	AA939876	AA939876	oC10d08.s	917	47	2.7	509	41	A1041011	A1041011	Dict
C 845	47	2.7	491	40	AA939876	AA939876	vz95h02.r	918	47	2.7	336	41	A1049659	A1049659	Dict
C 846	47	2.7	333	40	AA945339	AA945339	EST200838	919	47	2.7	71	41	A1049659	A1049659	Dict
C 847	47	2.7	127	40	AA946418	AA946418	EST201917	920	47	2.7	428	41	A1050084	A1050084	Dict
C 848	47	2.7	321	40	AA954676	AA954676	om95d10.s	921	47	2.7	423	41	A1051709	A1051709	Dict
C 849	47	2.7	345	40	AA958898	AA958898	ua19f03.r	922	47	2.7	445	41	A1056694	A1056694	Dict
C 850	47	2.7	437	40	AA959695	AA959695	vw56e06.s	923	47	2.7	69	41	A1061180	A1061180	Dict
C 851	47	2.7	343	40	AA960578	AA960578	vw64a12.s	924	47	2.7	279	41	A1061405	A1061405	Dict
C 852	47	2.7	391	40	AA969375	AA969375	om58e04.s	925	47	2.7	384	41	A1065658	A1065658	Dict
C 853	47	2.7	395	40	AA979714	AA979714	TC-EST-29	926	47	2.7	382	41	A1065658	A1065658	Dict
C 854	47	2.7	356	40	AA980543	AA980543	ua42e08.r	927	47	2.7	544	41	C93698	C93698	Dict
C 855	47	2.7	393	40	AA981912	AA981912	ua34c10.r	928	47	2.7	590	41	C93706	C93706	Dict
C 856	47	2.7	733	40	AA986840	AA986840	ue15f01.x	929	47	2.7	258	41	C93712	C93712	Dict
C 857	47	2.7	496	40	C89741	C89741	C89741.Dict	930	47	2.7	589	41	C93745	C93745	Dict
C 858	47	2.7	164	40	C89844	C89844	Dict	931	47	2.7	124	41	C93753	C93753	Dict
C 859	47	2.7	587	40	C89862	C89862	Dict	932	47	2.7	576	41	C93788	C93788	Dict
C 860	47	2.7	309	40	C89869	C89869	Dict	933	47	2.7	252	41	C93792	C93792	Dict
C 861	47	2.7	449	40	C89889	C89889	Dict	934	47	2.7	809	41	C93796	C93796	Dict

935	47	2.7	657	41	C93817	C93817	C93817	DICT
936	47	2.7	930	41	C93828	C93828	C93828	DICT
937	47	2.7	486	41	C93852	C93852	C93852	DICT
938	47	2.7	465	41	C93867	C93867	C93867	DICT
939	47	2.7	813	41	C93868	C93868	C93868	DICT
940	47	2.7	348	41	C93889	C93889	C93889	DICT
941	47	2.7	734	41	C93927	C93927	C93927	DICT
942	47	2.7	441	41	C93943	C93943	C93943	DICT
943	47	2.7	609	41	C93945	C93945	C93945	DICT
944	47	2.7	532	41	C94021	C94021	C94021	DICT
945	47	2.7	270	41	C94036	C94036	C94036	DICT
946	47	2.7	205	41	C94125	C94125	C94125	DICT
947	47	2.7	231	41	C94154	C94154	C94154	DICT
948	47	2.7	355	41	C94167	C94167	C94167	DICT
949	47	2.7	330	41	C94178	C94178	C94178	DICT
950	47	2.7	331	41	C94266	C94266	C94266	DICT
951	47	2.7	194	41	C94286	C94286	C94286	DICT
952	47	2.7	416	41	C94331	C94331	C94331	DICT
953	47	2.7	559	41	C94338	C94338	C94338	DICT
954	47	2.7	488	41	C94347	C94347	C94347	DICT
955	47	2.7	540	41	C94396	C94396	C94396	DICT
956	47	2.7	282	41	C94400	C94400	C94400	DICT
957	47	2.7	174	41	C94407	C94407	C94407	DICT
958	47	2.7	218	41	C94439	C94439	C94439	DICT
959	47	2.7	182	41	C94451	C94451	C94451	DICT
960	47	2.7	406	41	C94510	C94510	C94510	DICT
961	47	2.7	300	41	C94516	C94516	C94516	DICT
962	47	2.7	182	41	C95040	C95040	C95040	Pig
963	47	2.7	326	42	A1073383	A1073383	A1073383	0013d03.x
964	47	2.7	270	42	A1074030	A1074030	A1074030	0y66h06.x
965	47	2.7	400	42	A1075885	A1075885	A1075885	0z25d09.x
966	47	2.7	405	42	A1076157	A1076157	A1076157	0y92h04.x
967	47	2.7	636	42	A1079226	A1079226	A1079226	0z52b02.x
968	47	2.7	355	42	A1080256	A1080256	A1080256	0z47g06.x
969	47	2.7	291	42	A1081740	A1081740	A1081740	0u63g04.s
970	47	2.7	437	42	A1086783	A1086783	A1086783	0z77c02.x
971	47	2.7	458	42	A1088789	A1088789	A1088789	0z21c07.x
972	47	2.7	346	42	A1088929	A1088929	A1088929	0z21g02.x
973	47	2.7	279	42	A1089782	A1089782	A1089782	0z29d07.x
974	47	2.7	416	42	A1091468	A1091468	A1091468	0o19b03.x
975	47	2.7	361	42	A1095003	A1095003	A1095003	0z44d05.x
976	47	2.7	320	42	A1095192	A1095192	A1095192	0y83b09.s
977	47	2.7	712	42	A1096512	A1096512	A1096512	0z40c12.x
978	47	2.7	443	42	A1096694	A1096694	A1096694	0z81b08.x
979	47	2.7	342	42	A1097229	A1097229	A1097229	0z88d09.x
980	47	2.7	276	42	A1097410	A1097410	A1097410	0z44c10.x
981	47	2.7	91	42	A1105872	A1105872	A1105872	0z02a07.L
982	47	2.7	192	42	A1105948	A1105948	A1105948	0z03a04.L
983	47	2.7	223	42	A1105984	A1105984	A1105984	0z03f10.L
984	47	2.7	654	42	A1107087	A1107087	A1107087	0z06380.s
985	47	2.7	540	42	A1107657	A1107657	A1107657	0z05457.s
986	47	2.7	537	42	A1108100	A1108100	A1108100	0z06842.s
987	47	2.7	341	42	A1109273	A1109273	A1109273	0z08412.s
988	47	2.7	461	42	A1110228	A1110228	A1110228	0z052.ANY
989	47	2.7	530	42	A1110360	A1110360	A1110360	0z06685.s
990	47	2.7	205	42	A1115071	A1115071	A1115071	0z146e09.x
991	47	2.7	760	42	A1117098	A1117098	A1117098	0z03f06.Y
992	47	2.7	785	42	A1117621	A1117621	A1117621	0z148b03.x
993	47	2.7	564	42	A1125109	A1125109	A1125109	0z06f01.s
994	47	2.7	365	42	A1125884	A1125884	A1125884	0z06f04.x
995	47	2.7	492	42	A1128239	A1128239	A1128239	0z03f109.x
996	47	2.7	309	42	A1129698	A1129698	A1129698	0z035910.x
997	47	2.7	459	42	A1129719	A1129719	A1129719	0z04e009.x
998	47	2.7	795	42	A1133715	A1133715	A1133715	0z24e005.x
999	47	2.7	329	42	A1137715	A1137715	A1137715	0z03h05.x
1000	47	2.7	501	42	A1134068	A1134068	A1134068	0z11418.s

ALIGNMENTS

LOCUS	AA935648	589 bp	mRNA	EST	07-JUL-1998
DEFINITION	op27c09.s1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone IMAGE:1578064 3' similar to gb:L06797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCR1 HOMOLOG (HUMAN);, mRNA sequence.				
ACCESSION	AA935648				
NID	g3092805				
VERSION	AA935648.1 GI:3092805				
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
AUTHORS	1 (bases 1 to 589)				
TITLE	NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.				
JOURNAL	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index				
COMMENT	Unpublished (1997) On Jan 17, 1998 this sequence version replaced gi:1900670. Contact: Robert Strausberg, Ph.D. Tel: (301) 496-1550 Email: Robert_strausberg@nih.gov This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information. Insert Length: 1009 Std Error: 0.00 Seq primer: -40ml3 fwd. Ex from Amersham High quality sequence stop: 511. Location/Qualifiers 1..589 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:1578064" /clone_id="Soares_NFL_T_GBC_S1" /lab_host="DH10B" /note="Organ: pooled; Vector: pUT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not 1; Site_2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (fetal lung NBH19W, testis NT1, and B-cell NCI-CGAP GCBI) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bernaldo."				
FEATURES	Source				
BASE COUNT	220 a 116 c 89 g 164 t				
ORIGIN					
Query Match	33.9%; Score 589; DB 40; Length 589;				
Best Local Similarity	100.0%; Pred. No. 4,5e-245;				
Matches	589; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
OY	1087	AAGCAGGTGACATTCATCTGTTCCAGTGTGCTTCACGTTTCACGCCAC	1146		
DB	589	AAGCAGGTGACATTCATCTGTTCCAGTGTGCTTCACGTTTCACGCCAC	530		
OY	1147	TACACACATGTAAAGACTTTTTTTTACGATTAATACCTTTTTTAAAGTTACAT	1206		
DB	529	TACACACATGTAAAGACTTTTTTTTACGATTAATACCTTTTTTAAAGTTACAT	470		
OY	1207	TTTTCAGATTAATAAGACTGACCAATATTGATGTTTATGCTGTGGATTTTGT	1266		
DB	469	TTTTCAGATTAATAAGACTGACCAATATTGATGTTTATGCTGTGGATTTTGT	410		
OY	1267	CTTGTTCTCTTACGTTTGTGAAGTTAATGACTTATTAATAATTTTTTGT	1326		
DB	409	CTTGTTCTCTTACGTTTGTGAAGTTAATGACTTATTAATAATTTTTTGT	350		
OY	1327	TCATATTGATGTGTCTAGGAGAGACTGTGGCCAACTTCTTACTGCTGATCTCG	1386		
DB	349	TCATATTGATGTGTCTAGGAGAGACTGTGGCCAACTTCTTACTGCTGATCTCG	290		

QY 1387 TGTAGACACTAGAAAAAGCACTGAACATTCAGACGCTGTAGTGAATCAGCTAAAGC 1446
|||||
Db 289 TGTAGGAGCTGTAGAAAAAGGAGACGACATTCAGACGCTGTAGTGAATCAGCTAAAGC 230
QY 1447 TAGAATGATCCGACGCTGTATGATAGATATCTCCATTCGCGTGAAGCTTTT 1506
|||||
Db 229 TAGAAATGATCCGACGCTGTATGATAGATATCTCCATTCGCGTGAAGCTTTT 170
QY 1507 CCTGTCTTAAGACGTGATTTTGTCTGTAGAGATGACATTATACCAAGCCCAAGTG 1566
CCTGTCTTAAGACGTGATTTTGTCTGTAGAGATGACATTATACCAAGCCCAAGTG 110
QY 1567 GTATAGAAATGCTGTTTTCAGTTTTCAGAGTGGTGTGATTTAGACACCTACAGTGT 1626
|||||
Db 109 GTATAGAAATGCTGTTTTCAGTTTTCAGAGTGGTGTGATTTAGACACCTACAGTGT 50
QY 1627 CAGTCTGTATTAGTTGTATTAATAAGTACATGTTAACTTAAAAA 1675
CAGTCTGTATTAGTTGTATTAATAAGTACATGTTAACTTAAAAA 1
Db 49 CAGTCTGTATTAGTTGTATTAATAAGTACATGTTAACTTAAAAA 1

RESULT 2
AA411265 599 bp mRNA EST 17-MAY-1997
LOCUS z24h06.f1 Soares NhhpU.S1 Homo sapiens cDNA clone IMAGE:754619 5'
DEFINITION similar to gb:106797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCR1
HOMOLOG (HUMAN);, mRNA sequence.
ACCESSION AA411265
NID g2068847
VERSION AA411265.1 GI:2068847
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 599)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Giesel,G., Jost,S.,
Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J.,
Moore,B., Schellenberg,K., Stepien,M., Tan,F., Theising,B.,
White,Y., Wyllie,T., Waterston,R. and Wilson,R.
WashU-Merck EST Project 1997
TITLE Unpublished (1997)
JOURNAL On Sep 12, 1996 this sequence version replaced gi:1406946.
COMMENT

CONTACT: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 506.
Location/Qualifiers
1. 599
/organism="Homo sapiens"
/db_xref="GDB:5977529"
/db_xref="taxon:9606"
/clone_image="754619"
/clone_lib="Soares_NhhpU.S1"
/tissue_type="Pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="DH108"
/note="Organ: mixed (see below); Vector: pTT3-Pac
(Pharmacia) with a modified polylinker; Site:1; Noc I;
Site:2; Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2NblM, pregnant uterus
NblPU, and fetal heart NblH19M) were mixed, and ss circles
were made in vitro. Following HAP purification, this DNA
was used as tracer in a subtractive hybridization
reaction. The driver was PCR amplified cDNAs from pools of

5,000 clones made from the same 3 libraries. The pools
consisted of T.M.A.G.E. clones 250232-265223,
340488-345479, and 484468-489479."

BASE COUNT 168 a 93 c 117 g 221 t
ORIGIN

Query Match 32.3%; Score 561; DB 33; Length 599;
Best Local Similarity 100.0%; Pred. No. 5.9e-233;
Matches 561; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1092 AGGTGACATTCATGTTTTCACATGATGAGCTTCAGATTTTCACTCCAGCTAC 1151
|||||
Db 15 AGGTGACATTCATGTTTTCACATGATGAGCTTCAGATTTTCACTCCAGCTAC 74
QY 1152 CAGATGTAAGACACTTTTATACGATTAATACATTTTATAGTACATTTTC 1211
|||||
Db 75 CAGATGTAAGACACTTTTATACGATTAATACATTTTATAGTACATTTTC 134
QY 1212 AGATATTAAGACATGACCAATATGTTACAGTTTATGCTTGTGATTTTGTCTGT 1271
AGATATTAAGACATGACCAATATGTTACAGTTTATGCTTGTGATTTTGTCTGT 194
Db 135 AGATATTAAGACATGACCAATATGTTACAGTTTATGCTTGTGATTTTGTCTGT 194
QY 1272 GTTCTTGTAGTTTGTGAGTTTATGACTTATGATTAATATTTTGTTCATA 1331
|||||
Db 195 GTTCTTGTAGTTTGTGAGTTTATGACTTATGATTAATATTTTGTTCATA 254
QY 1332 TTGATGTGTCTAGCAGACCTGTGCGCAAGTTCTAGTGTGATGCTGTGTA 1391
|||||
Db 255 TTGATGTGTCTAGCAGACCTGTGCGCAAGTTCTAGTGTGATGCTGTGTA 314
QY 1392 GGACGTGTAAGAAAGGACCTGACATTCACAGCCTGTAGTGAATCAGCTAAGCTAGAA 1451
GGACGTGTAAGAAAGGACCTGACATTCACAGCCTGTAGTGAATCAGCTAAGCTAGAA 374
Db 315 GGACGTGTAAGAAAGGACCTGACATTCACAGCCTGTAGTGAATCAGCTAAGCTAGAA 374
QY 1452 ATGATCCCGACGCTGTTTATGATAGATATCTCCATTCGCGTGAAGCTTTTCTGT 1511
ATGATCCCGACGCTGTTTATGATAGATATCTCCATTCGCGTGAAGCTTTTCTGT 434
Db 375 ATGATCCCGACGCTGTTTATGATAGATATCTCCATTCGCGTGAAGCTTTTCTGT 434
QY 1512 TCTTAAGACGTGATTTTCTGTAGAGATGACCTTATACCAAGCCCAAGTGTATA 1571
TCTTAAGACGTGATTTTCTGTAGAGATGACCTTATACCAAGCCCAAGTGTATA 494
Db 435 TCTTAAGACGTGATTTTCTGTAGAGATGACCTTATACCAAGCCCAAGTGTATA 494
QY 1572 GAAATGCTGTTTTCAGTTTTCAGAGTGGTGTGATTTTCAGACCTACAGTACATTC 1631
GAAATGCTGTTTTCAGTTTTCAGAGTGGTGTGATTTTCAGACCTACAGTACATTC 554
Db 495 GAAATGCTGTTTTCAGTTTTCAGAGTGGTGTGATTTTCAGACCTACAGTACATTC 554
QY 1632 TTGTATTAGTTGTTAATATA 1652
TTGTATTAGTTGTTAATATA 575
Db 555 TTGTATTAGTTGTTAATATA 575

RESULT 3
AA610463/c 539 bp mRNA EST 30-OCT-1997
LOCUS np97b09.s1 NCI-CGAP Lul Homo sapiens cDNA clone IMAGE:1142297 3'
DEFINITION similar to gb:106797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCR1
HOMOLOG (HUMAN);, mRNA sequence.
ACCESSION AA610463
NID g2458891
VERSION AA610463.1 GI:2458891
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 539)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1395020.

Db 338 TTGATGTGTCTAGCAGACACTGTGGCAAGTCTTAGTTCGTATGTCGTGTA 279
QY 1392 GGACGTGTACAAAAAGGAACTGAACATTCAGACCGGTGTACTGATCATCGTAAAGCTAGAA 1451
Db 278 GGACGTGTAGAAAAAGGAACTGAACATTCAGACCGGTGTACTGATCATCGTAAAGCTAGAA 219
QY 1452 ATGATCCCGCAGCGTGTATGATGATGAATACCTCCATTCCTCGGTGAAGCTTTTCTGT 1511
Db 218 ATGATCCCGCAGCGTGTATGATGATGAATACCTCCATTCCTCGGTGAAGCTTTTCTGT 159
QY 1512 TCTTAAGACGTATTTGCTGTAGAGATGGACCTATTAACCAAGCCCAAGTGTATA 1571
Db 158 TCTTAAGACGTATTTGCTGTAGAGATGGACCTATTAACCAAGCCCAAGTGTATA 99
QY 1572 GAAATGCTGCTTTTCAGTTTTCAGAGATGGGTGTATTTTCAGACCTACAGTGTACAGTC 1531
Db 98 GAAATGCTGCTTTTCAGTTTTCAGAGATGGGTGTATTTTCAGACCTACAGTGTACAGTC 39
QY 1532 TTGTATTAGTGTATTAAGATGATGATTAACCTTA 1669
Db 38 TTGTATTAGTGTATTAAGATGATGATTAACCTTA 1
RESULT 5
AA426644/C 501 bp mRNA EST 16-OCT-1997
LOCUS zva7h11.s1 Soares ovary tumor NBHOT Homo sapiens cDNA clone
DEFINITION IMAGE:756837.3; similar to gb:106797 PROBABLE G PROTEIN-COUPLED
RECEPTOR LCRI HOMOLOG (HUMAN);, mRNA sequence.
ACCESSION AA426644
NID 92107314
VERSION AA426644.1 GI:2107314
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 501)
Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisels,G., Jost,S.,
Kucaba,T., Lacey,M., Le,N., Lennon,G., Matra,M., Martin,J.,
Moore,B., Scheilenberg,K., Steptoe,M., Tan,F., Theising,B.,
White,Y., Wyllie,T., Waterston,R. and Wilson,R.
WashU-Merck EST Project 1997
Unpublished (1997)
On Apr 14, 1993 this sequence version replaced gi:693425.
TITLE
JOURNAL
COMMENT
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -41m13 fwd. Et from Amersham
High quality sequence stop: 332.
Location/Qualifiers
1..501
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="7"
/clone="IMAGE:756837"
/clone_lib="Soares ovary tumor NBHOT"
/sex="Female"
/tissue_type="ovarian tumor"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: ovary; Vector: pRT3D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer (5'
TGTTACCAATCTGAAGTGGAGCGCGCGGTTTCTTTTCTTTT 3'),
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pRT3D vector

(Pharmacia), library constructed by Bento Soares and
M.Fatima Bonadio.
BASE COUNT 185 a 99 c 69 g 148 t
ORIGIN
Query Match 28.8%; Score 501; DB 33; Length 501;
Best Local Similarity 100.0%; Pred. No. 5.8e-20;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1190 TTTTAACTTACACATTTTTCAGATATTAAGACGTACCAATATTTGACGTTTATT 1249
Db 501 TTTTAACTTACACATTTTTCAGATATTAAGACGTACCAATATTTGACGTTTATT 442
QY 1250 GCTTGTGCAATTTTGTCTTGTCTTCTTATTTTGTGAACTTAATTTGACTTATTA 1309
Db 441 GCTTGTGCAATTTTGTCTTGTCTTCTTATTTTGTGAACTTAATTTGACTTATTA 382
QY 1310 TATTAATTTTCTTGTCTTATATGATGCTGTCTAGGAGGACCTGTGGCCAGTTCTT 1369
Db 381 TATTAATTTTCTTGTCTTATATGATGCTGTCTAGGAGGACCTGTGGCCAGTTCTT 322
QY 1370 AGTTGCTGATGCTGTCTGTAGAGCTGTAAGAAAGGAACTGAACTTCCAGAGCGTGT 1429
Db 321 AGTTGCTGATGCTGTCTGTAGAGCTGTAAGAAAGGAACTGAACTTCCAGAGCGTGT 262
QY 1430 AGTGAATCACTGAACCTGAAATGATCCCGACCTTTTATGATAGATATCTCCAT 1489
Db 261 AGTGAATCACTGAACCTGAAATGATCCCGACCTTTTATGATAGATATCTCCAT 202
QY 1490 TCCCGTGAAGCGTTTTCCTGCTTAAGACGTATTTGCTGTAGAGATGGCACTTAT 1549
Db 201 TCCCGTGAAGCGTTTTCCTGCTTAAGACGTATTTGCTGTAGAGATGGCACTTAT 142
QY 1550 AACCAAGCCCAAGTGTATGAAATGCTGTTTTCAGTTTTCAGAGTGGTGTGATT 1609
Db 141 AACCAAGCCCAAGTGTATGAAATGCTGTTTTCAGTTTTCAGAGTGGTGTGATT 82
QY 1610 TCAGCACTGACGTGACGTGCTGTATTAAGTGTATTAAGATGATGTTAAACTTA 1669
Db 81 TCAGCACTGACGTGACGTGCTGTATTAAGTGTATTAAGATGATGTTAAACTTA 22
QY 1670 AAAAAAAAAAAAAAAAAAAAAA 1690
Db 21 AAAAAAAAAAAAAAAAAAAAAA 1
RESULT 6
A1493618 657 bp mRNA EST 30-MAR-1999
LOCUS A1493618/C
DEFINITION th39e07.x1 NCI-CGAP Pauli Homo sapiens cDNA clone IMAGE:2120676.3;
similar to gb:106797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCRI
HOMOLOG (HUMAN);, mRNA sequence.
ACCESSION A1493618
NID 94394621
VERSION A1493618.1 GI:4394621
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 657)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On May 18, 1998 this sequence version replaced gi:3137281.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Life Technologies catalog #: 11548-013
DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bdrrp/image/image.html

Insert Length: 1615 Std Error: 0.00
Seq primer: -40UP from Gldco
High quality sequence stop: 410.

FEATURES

SOURCE

1. .657
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2120676"
/clone_id="NCI CGAP_Pan1"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/note="Organ: pancreas; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.72 kb. Life Technologies catalog #: 1548-013"

BASE COUNT 232 a 130 c 109 g 184 t 2 others
ORIGIN

Query Match 28.7%; Score 498; DB 47; Length 657;
Best Local Similarity 99.8%; Pred. No. 1.1e-205;
Matches 548; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1129 TCAAGTTTCACTCCAGCTAACACAGATGTAAAGACTTTTATACGATAAATACT 1188
DB 549 TCAAGTTTCACTCCAGCTAACACAGATGTAAAGACTTTTATACGATAAATACT 490
QY 1189 TTTTATAGTACACATTTTTCAGATATAAAGACTGCCAATATTGACGTTTAT 1248
DB 489 TTTTATAGTACACATTTTTCAGATATAAAGACTGCCAATATTGACGTTTAT 430
QY 1249 TGCTTGTCGATTTTGTCTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1308
DB 429 TGCTTGTCGATTTTGTCTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 370
QY 1309 ATATATAATTTTTTTTGTTCATATTGATGTGTCTAGGACGACCTGTGCCAAGTTCT 1368
DB 369 ATATATAATTTTTTTTGTTCATATTGATGTGTCTAGGACGACCTGTGCCAAGTTCT 310
QY 1369 TAGTGTCTGATGCTCGGGGTAGACGTAGAAAAGGAACTGAAACATTCAGACGCG 1428
DB 309 TAGTGTCTGATGCTCGGGGTAGACGTAGAAAAGGAACTGAAACATTCAGACGCG 250
QY 1429 TAGTGAATCAGCTAAGCTAGAAATGATCCCGAGCTGTTTATGATAGATAATCTCTCA 1488
DB 249 TAGTGAATCAGCTAAGCTAGAAATGATCCCGAGCTGTTTATGATAGATAATCTCTCA 190
QY 1489 TTCCCGTGGAGCTTTTCTCTTCTTAAAGCTGATTTTGTCTGTAAGAATGCGACTTA 1548
DB 189 TTCCCGTGGAGCTTTTCTCTTCTTAAAGCTGATTTTGTCTGTAAGAATGCGACTTA 130
QY 1549 TAACCAAGCCCAAGGTGTATAGAAATCTGGTTTTCAGTTTCAGAGAGGGGTGAT 1608
DB 129 TAACCAAGCCCAAGGTGTATAGAAATCTGGTTTTCAGTTTCAGAGAGGGGTGAT 70
QY 1609 TTCACACCTACAGTACAGTCTGTATTAAGTTGTTAATAAAGTACATGTTAACTT 1668
DB 69 TTCACACCTACAGTACAGTCTGTATTAAGTTGTTAATAAAGTACATGTTAACTT 10
QY 1669 AAAAAAAAAA 1677
DB 9 AAAAAAAAAA 1

RESULT 7
AA634211/c 693 bp mRNA EST 06-MAR-1998
LOCUS AA634211/c
DEFINITION ac72b06.s1 Strategene Lung (#937210) Homo sapiens CDNA clone
IMAGE:868115 3' similar to gb:106797 PROBABLE G PROTEIN-COUPLED

ACCESSION AA634211
NID 02557425
VERSION AA634211.1 GI:2557425
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 693)
AUTHORS Hallier, U., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Kitzman, D., Kuaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wyllie, T., Waterston, R. and Wilson, R.
TITLE WashU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1393895.

CONTACT: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu

This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 835 Std Error: 0.00
Seq primer: -40m13 fwd. RT from Amersham
High quality sequence stop: 393.
Location/Qualifiers

FEATURES

SOURCE

1. .693
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:868115"
/clone_id="Stratagene Lung (#937210)"
/sex="male"
/dev_stage="72 years"
/lab_host="SOLR cells (kanamycin resistant)"
/note="Organ: lung; Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dt. normal lung. Average insert size: 1.0 kb; Uni-ZAP XR Vector: -5' adaptor sequence: 5' GAATTCGGCAGC 3' -3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTT 3'"

Query Match 28.0%; Score 486; DB 36; Length 693;
Best Local Similarity 99.8%; Pred. No. 1.7e-200;
Matches 536; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1135 TTTCACCTCCAGTACACAGATGTAAAGACTTTTATACGATAAATACTTTT 1194
DB 537 TTTCACCTCCAGTACACAGATGTAAAGACTTTTATACGATAAATACTTTTCT 478
QY 1195 TAAGTTACACATTTTTCAGATATAAAGACTGACCAATATTGACGTTTATGCTTG 1254
DB 477 TAAGTTACACATTTTTCAGATATAAAGACTGACCAATATTGACGTTTATGCTTG 418
QY 1255 TTGATTTTGTCTGTGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1314
DB 417 TTGATTTTGTCTGTGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 358
QY 1315 ATTTTCTTGTTCATATTGATGTGTCTAGGACGACCTGTGCCAAGTCTTAGTTG 1374
DB 357 ATTTTCTTGTTCATATTGATGTGTCTAGGACGACCTGTGCCAAGTCTTAGTTG 298
QY 1375 CTGTATGCTCGTGTAGAGCTGTAGAAAGGAGACTGACATTCACAGCGCTAGTGA 1434
DB 297 CTGTATGCTCGTGTAGAGCTGTAGAAAGGAGACTGACATTCACAGCGCTAGTGA 238
QY 1435 ATCAGCTAAAGCTAGAAATGATCCCGAGCTTTATGATAGATAATCTTCATTCGCG 1494

Db 237 ATCAGCTAAGCTAGAAATGATGCCAGCTGTTATGATGATAGATATCTCTCCATTCGG 178

QY 1495 TGGACGTTTTCCTGTTCTTTAGACGTATTTGCTGTAGAAATGGACCTATATACCA 1554

Db 177 TGGACGTTTTCCTGTTCTTTAGACGTATTTGCTGTAGAAATGGACCTATATACCA 118

QY 1555 AAGCCCAAGTGTATAGAAATGCTGTTTTCAGTTTTCAGAGATGGGTGATTTAGC 1614

Db 117 AAGCCCAAGTGTATAGAAATGCTGTTTTCAGTTTTCAGAGATGGGTGATTTAGC 58

QY 1615 ACCTACAGTGTACAGTCTGTATTAAGTTGTATATAAAGTACATGTTAACTTAA 1671

Db 57 ACCTACAGTGTACAGTCTGTATTAAGTTGTATATAAAGTACATGTTAACTTAA 1

RESULT 8
AA634201/c 651 bp mRNA EST 21-OCT-1997
LOCUS ac27a07.s1 Stratagene lung (#937210) Homo sapiens cDNA clone
DEFINITION IMAGE:868116.3, similar to gb:L06797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCRI HOMOLOG (HUMAN);, mRNA sequence.

ACCESSION AA634201
NID 92557415
VERSION AA634201.1 GI:2557415
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisels,G., Jost,S., Martin,D., Kucaba,T., Lacy,M., Lennon,G., Marra,M., Marini,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R. WashU-MCI human EST Project
TITLE Unpublished (1997)
JOURNAL On Sep 12, 1996 this sequence version replaced gi:1393737.
COMMENT

CONTACT: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL: contact the IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: -40m13 fwd. Et from Amersham
High quality sequence stop: 413.

FEATURES
source
1..651
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:868116"
/clone_lib="Stratagene lung (#937210)"
/sex="male"
/dev_stage="72 years"
/lab_host="SOLR cells (kanamycin resistant)"
/note="Organ: lung; Vector: pBluescript SK-; Site:1: EcoRI; Site-2: XhoI; Cloned unidirectionally. Primer: Oligo df. normal lung; Average insert size: 1.0 kb; Uni-ZAP XR Vector: -5' adaptor sequence: 5' GAATTCGGCAGAG 3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'."
BASE COUNT 230 a 130 c 112 g 179 t
ORIGIN

Query Match 27.6%; Score 480; DB 36; Length 651;
Best Local Similarity 100.0%; Pred. NO. 6.7e-198;
Matches 480; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1192 TTTTAGTACATTTTCAGATATATAAGCTGACCAATATGTACGTTTTCG 1251
|||||
Db 480 TTTTAGTACATTTTCAGATATATAAGCTGACCAATATGTACGTTTTCG 421

QY 1252 TGTGATTTTTTGTCTGTTGTTCTTTAGTTTTCGAGTTTAATGACTATTATTA 1311
|||||
Db 420 TGTGATTTTTTGTCTGTTGTTCTTTAGTTTTCGAGTTTAATGACTATTATTA 361

QY 1312 TAAATTTTTTTTGTTCATATATGATGTGTCTAGCAGACCTGTGGCCAATGTTAG 1371
|||||
Db 360 TAAATTTTTTTTGTTCATATATGATGTGTCTAGCAGACCTGTGGCCAATGTTAG 301

QY 1372 TTGCTGATGTCTGCTGTGTGAGACTGTAGAAAGGAAAGCAATCCAGAGGTAG 1431
|||||
Db 300 TTGCTGATGTCTGCTGTGTGAGACTGTAGAAAGGAAAGCAATCCAGAGGTAG 241

QY 1432 TGAATCAGTAAAGCTAGAAATGATCCCGCTGTTTATGATAGATATCTCCATTC 1491
Db 240 TGAATCAGTAAAGCTAGAAATGATCCCGCTGTTTATGATAGATATCTCCATTC 181

QY 1492 CGGTGAGCTTTTCTGTTCTTTAGAGCTGATTTTGTGTAGAGATGCACTTAA 1551
Db 180 CGGTGAGCTTTTCTGTTCTTTAGAGCTGATTTTGTGTAGAGATGCACTTAA 121

QY 1552 CCAAGGCCAAGTGTATAGAAATGCTGTTTTCAGTTTTCAGAGATGGTTGATTC 1611
|||||
Db 120 CCAAGGCCAAGTGTATAGAAATGCTGTTTTCAGTTTTCAGAGATGGTTGATTC 61

QY 1612 AGCAGCTACAGTGTACAGTCTGTATTAAGTTGTATATAAAGTACATGTTAACTTAA 1671
Db 60 AGCAGCTACAGTGTACAGTCTGTATTAAGTTGTATATAAAGTACATGTTAACTTAA 1

RESULT 9
AA426566 470 bp mRNA EST 16-OCT-1997
LOCUS z47h11.r1 Soares ovary tumor NBOT Homo sapiens cDNA clone
DEFINITION IMAGE:756837.5, similar to gb:L06797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCRI HOMOLOG (HUMAN);, mRNA sequence.

ACCESSION AA426566
NID 92107387
VERSION AA426566.1 GI:2107387
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisels,G., Jost,S., Kucaba,T., Lacy,M., Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R. WashU-Merck EST Project 1997
TITLE Unpublished (1997)
JOURNAL On Sep 12, 1996 this sequence version replaced gi:1288644.
COMMENT

CONTACT: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL: contact the IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: -28m13 rev2 Et from Amersham
High quality sequence stop: 287.

FEATURES
source
1..470
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="5; 6q25-26"
/clone="IMAGE:756837"
/clone_lib="Soares ovary tumor NBOT"
/sex="Female"
/tissue_type="ovarian tumor"
/lab_host="DH10B (ampicillin resistant)"

/note="Organ: ovary; Vector: pT73D (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA was primed with a Not I...oligo(61) primer [5' TGTTCACATCTGGAAGTGGAGCGCGGGGTTTGTGTTTGTGTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library constructed by Bento Soares and M. Fatima Bernaldo."

BASE COUNT 134 a 66 c 98 g 172 t
ORIGIN

Query Match 27.1% Score 470; DB 33; Length 470;
Best Local Similarity 100.0%; Pred. No. 1.5e-193;
Matches 470; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1213 GATATTAAGAGTACCAATATTTGACAGTTTATGCTGTGATTTTGTCTGTG 1272
DB 1 GATATTAAGAGTACCAATATTTGACAGTTTATGCTGTGATTTTGTCTGTG 60
OY 1273 TTTCTTACTTTTGTGAGCTTTAATGACTTTATATATAATTTTGTCTGTGAT 1332
DB 61 TTTCTTACTTTTGTGAGCTTTAATGACTTTATATATAATTTTGTCTGTGAT 120
OY 1333 TGATGTGTCTAGGAGGAGCCTGTGGCCAGTTCTTGTGTGTGTGTGTGTG 1392
DB 121 TGATGTGTCTAGGAGGAGCCTGTGGCCAGTTCTTGTGTGTGTGTGTGTG 180
OY 1393 GACTGTAGAAAAGGAGACATTCAGAGCGTGTAGTAATCAGTAAAGCTAGAAA 1452
DB 181 GACTGTAGAAAAGGAGACATTCAGAGCGTGTAGTAATCAGTAAAGCTAGAAA 240
OY 1453 TGATCCCCCGCTGTATGATATATCTCCATTCGCCGGAAGCTTTTCTGTG 1512
DB 241 TGATCCCCCGCTGTATGATATATCTCCATTCGCCGGAAGCTTTTCTGTG 300
OY 1513 CTTAAGACGCTGATTTTGTGTAGAGATGCGACTTATTAACCAAGCCCAAGTGTATAG 1572
DB 301 CTTAAGACGCTGATTTTGTGTAGAGATGCGACTTATTAACCAAGCCCAAGTGTATAG 360
OY 1573 AATGCTGCTTTTCACTTTTCAAGAGTGGGTTGATTTAGCAGCCTACAGTGTACGCT 1632
DB 361 AATGCTGCTTTTCACTTTTCAAGAGTGGGTTGATTTAGCAGCCTACAGTGTACGCT 420
OY 1533 TCTATTAACTGTTAATAAAGTCACTTTAACTTAAAAA 1682
DB 421 TCTATTAACTGTTAATAAAGTCACTTTAACTTAAAAA 470

RESULT 10
AA804282/C
LOCUS AA804282 536 bp mRNA EST 18-FEB-1998
DEFINITION n30308.s1 NCI_CGAP GC80 Homo sapiens cDNA clone IMAGE:142011 3'
similar to gb:L06797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCRI
HOMOLOG (HUMAN); mRNA sequence.

ACCESSION AA804282
NID 92873569
VERSION AA804282.1 GI:2873569
KEYWORDS EST
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 536)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
JOURNAL Tumor Gene Index
COMMENT Unpublished (1997)
On Jan 14, 1998 this sequence version replaced gi:1798484.

CONTACT: Robert Strausberg, Ph.D.
Tel.: (301) 496-1350

Email: Robert.Strausberg@nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman, Ph.D., Gerald Marti, M.D.
cDNA Library Preparation: Louis M. Staudt, M.D., Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/dbp/image/image.html

FEATURES
source
Insert length: 1513 Std Error: 0.00
Seq primer: -40ml3 fwd. Et from Amersham
High quality sequence stop: 486.
Location/Qualifiers
1..536
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="21"
/clone="IMAGE:1242011"
/clone_id="NCI_CGAP_GC80"
/tissue_type="germinal center B-cells"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: tonsil; Vector: Bluescript SK-; Site.1: EcoRI; Site.2: XhoI; Cloned unidirectionally. Primer: Oligo dt. Germinal center B-cells Library constructed by Dr. L. Staudt (NCI). 5' adaptor sequence: 5' GAATTCGCGCAGAG 3' 3' adaptor sequence: 5' CTCGAGTCTTTTGTGTTTGTGTTT 3' Average insert size: 1.1 kb."

BASE COUNT 195 a 107 c 81 g 153 t
ORIGIN

Query Match 27.0% Score 469; DB 38; Length 536;
Best Local Similarity 100.0%; Pred. No. 4e-193;
Matches 469; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1210 TCAGATATTAAGAGTACCAATATTTGACAGTTTATGCTGTGATTTTGTCTT 1269
DB 469 TCAGATATTAAGAGTACCAATATTTGACAGTTTATGCTGTGATTTTGTCTT 410
OY 1270 GTGTTTCTTACTTTTGTGAAAGTTAATGACTTATATATAATTTTGTGTTCA 1329
DB 409 GTGTTTCTTACTTTTGTGAAAGTTAATGACTTATATATAATTTTGTGTTCA 350
OY 1330 TATTGATGTGTCTAGGAGGAGCCTGTGGCCAGTTCTTGTGTGTGTGTGTGTG 1389
DB 349 TATTGATGTGTCTAGGAGGAGCCTGTGGCCAGTTCTTGTGTGTGTGTGTGTG 290
OY 1390 TAGGACTGTAGAAAAGGAGAACTGAACTTCAGAGCGTGTAGTAATCAGTAAAGCTAG 1449
DB 289 TAGGACTGTAGAAAAGGAGAACTGAACTTCAGAGCGTGTAGTAATCAGTAAAGCTAG 230
OY 1450 AATGATCCCCCGCTGTATGATATATCTCTCCATTCGCCGGAAGCTTTTCT 1509
DB 229 AATGATCCCCCGCTGTATGATATATCTCTCCATTCGCCGGAAGCTTTTCT 170
OY 1510 GTTCTTAAGAGCTGATTTTGTGTAGAGATGGCACTTATAACCAAGCCCAAGTGTGA 1569
DB 169 GTTCTTAAGAGCTGATTTTGTGTAGAGATGGCACTTATAACCAAGCCCAAGTGTGA 110
OY 1570 TAGAATGCTGTTTTCAGATTTTCAGAGTGGGTTGATTTTCAGACCTACAGTGTACAG 1629
DB 109 TAGAATGCTGTTTTCAGATTTTCAGAGTGGGTTGATTTTCAGACCTACAGTGTACAG 50
OY 1630 TCTTGTATTAACTGTTAATAAAGTCACTTTAACTTAAAAA 1678
DB 49 TCTTGTATTAACTGTTAATAAAGTCACTTTAACTTAAAAA 1

RESULT 11
AA781110/C
LOCUS AA781110 792 bp mRNA EST 31-DEC-1998
DEFINITION aJ3e10.s1 Soares_testis_NHT Homo sapiens cDNA clone J391178 3'

similar to gb:L06797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCRI
HOMOLOG (HUMAN);, mRNA sequence.

ACCESSION
AA781110
NID
92840441
VERSION
AA781110.1 GI:2840441
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens

REFERENCE
Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Mammalia:
Eutheria: Primates: Catarrhini: Homidae: Homo.
1 (bases 1 to 792)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
On Sep 12, 1996 this sequence version replaced gi:1398066.

JOURNAL
COMMENT

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bdrip/image/image.html

Insert Length: 1123 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 474.

FEATURES
Source
Location/Qualifiers

1..792
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="1391178"
/clone_lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand CDNA
was prepared from mRNA obtained from Clontech
Laboratories, Inc., and primed with a Not I - oligo(dT)
primer [5'
IGTTACCACTCGAGTGGAGCGGCCCAATTTTCTTTTCTT 3']
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 264 a 160 c 154 g 214 t
ORIGIN

Query Match 26.6% Score 462 DB 38 Length 792
Best Local Similarity 100.0% Pred. No. 3.9e-190
Matches 462: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

OY 1210 TCAGATATAAAGACTGACCAATATTTGTACAGTTTATTCCTGTGGATTTCCTT 1269
|||||
DB 462 TCAGATATAAAGACTGACCAATATTTGTACAGTTTATTCCTGTGGATTTCCTT 403
OY 1270 GTGTTCTTACATTTTGTGAGAGTTAATTGACTTATTAATTTTGTGTTCA 1329
|||||
DB 402 GTGTTCTTACATTTTGTGAGAGTTAATTGACTTATTAATTTTGTGTTCA 343
OY 1330 TATGATGTGTCTAGCAGGAGACCTGTGGCCAACTTCTAGTGTCTGTCTG 1389
|||||
DB 342 TATGATGTGTCTAGCAGGAGACCTGTGGCCAACTTCTAGTGTCTGTCTG 283
OY 1390 TAGGACTGTAGAAAGGAGACTGACATTCACAGCGGTAGAGATCAGTAACCTAG 1449
|||||
DB 282 TAGGACTGTAGAAAGGAGACTGACATTCACAGCGGTAGAGATCAGTAACCTAG 223

OY 1450 AAATGATCCCACTGTTTATGCAATATATCTTCACATTCGCCGTGGAAGCTTTTCTT 1509
|||||
DB 222 AAATGATCCCACTGTTTATGCAATATATCTTCACATTCGCCGTGGAAGCTTTTCTT 163
OY 1510 GTTCTTAGACGATGATTTTCTGATAGAGATGACCTTATACCAAGCCCAAGTGTGA 1569
|||||
DB 162 GTTCTTAGACGATGATTTTCTGATAGAGATGACCTTATACCAAGCCCAAGTGTGA 103
OY 1570 TAGAATGCTGCTTTTTCAGTTTTCAGAGAGTGGTGTGATTTAGACACCTACAGTACAG 1629
|||||
DB 102 TAGAATGCTGCTTTTTCAGTTTTCAGAGAGTGGTGTGATTTAGACACCTACAGTACAG 43
OY 1630 TCTTGTATTAGTTGTTAATAAAGTACATGTTAACTTAA 1671
|||||
DB 42 TCTTGTATTAGTTGTTAATAAAGTACATGTTAACTTAA 1

RESULT 12
AA743645/C 556 bp mRNA EST 22-JAN-1998
LOCUS
DEFINITION
ny24g09.s1 NCI-CGAP GCBI Homo sapiens cDNA clone IMAGE:1272736 3'
similar to gb:L06797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCRI
HOMOLOG (HUMAN);, mRNA sequence.

ACCESSION
NID
92783151
VERSION
AA743645.1 GI:2783151
KEYWORDS
SOURCE
human.
ORGANISM

REFERENCE
Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Mammalia:
Eutheria: Primates: Catarrhini: Homidae: Homo.
1 (bases 1 to 556)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
On Jan 9, 1998 this sequence version replaced gi:949470.

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
Ph.D., Gerald Marti, M.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bdrip/image/image.html

Insert Length: 742 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 389.

FEATURES
Source
Location/Qualifiers

1..556
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1272736"
/clone_lib="NCI-CGAP GCBI"
/tissue="germinal center B cell"
/lab_host="DH10B"
/note="vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand CDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, IgD-),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGAAGTGGAGCGGCCCTCATTTTCTTTTCTTTT-
3']. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I

Page 16

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html

```

FEATURES
  source          1..455
                    location/Qualifiers
High quality sequence stop: 401.
Seq primer: -40up from GIBCO
Insert Length: 1080      Std Error: 0.00

```

1.455

double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I

BASE COUNT	168 a	96 c	65 g	126 t
ORIGIN	constructed by Bento Soares and M. Fatima Bonaldo. "			

Query Match	25.4%	Score 441	DB 44	Length 455
Best Local Similarity	100.0%	Pred. No. 5.5e+18		
Matches 441	Conservative	0	Mismatches	0
		Indels		0

[illegible]

Accession	Sequence	Position
QY	ACGTTTATTAATTTTTTGTTCATATGATGTGTCTGACGAGACCTGTGTGC	1350
Db	ACGTTTATTAATTTTTTGTTCATATGATGTGTCTGACGAGACCTGTGTGC	322
QY	ACGTTTATTAATTTTTTGTTCATATGATGTGTCTGACGAGACCTGTGTGC	1350
Db	ACGTTTATTAATTTTTTGTTCATATGATGTGTCTGACGAGACCTGTGTGC	322

Accession	Sequence	Length
Db	321 CAAGTTCTTAATGCTGTATGCTCGGTGAGACGTGAGAAAGGAACGAAACATTC	262
OY	1421 AGAGCGTGTACTGATACGTAAGCTACGAATGATCCCAAGCTGTTATGATAGATAA	1480

Db 261 AGAGCGGTAGTGAATACGTAAGGTAGAAATGATCCCACTGTTATCATAGATAA 202

1481 TCTCTCCATTCCGCTGGAACGTTTTTCCCTGTTCTTAAGACGTGATTTTGCATGAAGAT 1540

1541 GGCACCTATAACCAACCAACCCAAAGTGTGTAAAAAATGCTGTTTTCAGTTTCAGAGAGT 1600
141 GGCACCTATAACCAACCAACCCAAAGTGTGTAAAAAATGCTGTTTTCAGTTTTCAGAGAGT 93

Qy 1601 GGGTGGATTTCAGACCTACAGTGCACACTCTTTATTAACTGTAATAAAAGTACATG 1660
|||||
Db 81 GGGTGGATTTCAGACCTACAGTGCACACTCTTTATTAACTGTAATAATAAAAGTACATG 22

[illegible]

RESULT 14
AI344724/c
LOCUS
AI344724 463 bp mRNA
EST
02-FEB-1999

ACCESSION
ID
94081930
AI344724
similar to gb:L06797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCRI
HOMOLOG (HUMAN); mRNA sequence.


```
|||||
Db 437 CAGTTTATGCTGTGATTTTCTGTGTTCTTAGTTTGTGAAGTTTAA 378
Oy 1299 TGACTATATTAATTTTTTTTGTTCATATGATGTGTAGGAGACCTGTG 1358
Db 377 TGACTATATTAATTTTTTTTGTTCATATGATGTGTAGGAGACCTGTG 318
Oy 1359 GCGAAGTTCTAGTGTGATGTCTGTGTGTAGAGACTAGAAAAGGAACTGAACATT 1418
Db 317 GCGAAGTTCTAGTGTGATGTCTGTGTGTAGAGACTAGAAAAGGAACTGAACATT 258
Oy 1419 CCAGACGCTAGTGAATCACTAAGCTAAGAAATGATCCCGAGCTGTTATGCATGAT 1478
Db 257 CCAGACGCTAGTGAATCACTAAGCTAAGAAATGATCCCGAGCTGTTATGCATGAT 198
Oy 1479 AATCTCTCCATCCCGGAGAACGTTTTTCTGTCTTAAACGCGATTTTGTCTGTAAG 1538
Db 197 AATCTCTCCATCCCGGAGAACGTTTTTCTGTCTTAAACGCGATTTTGTCTGTAAG 138
Oy 1539 ATGCACTTTAACCAGCCCAAGTGTATAGAAATGCTGTTTTCAGTTTTCAGGA 1598
Db 137 ATGCACTTTAACCAGCCCAAGTGTATAGAAATGCTGTTTTCAGTTTTCAGGA 78
Oy 1599 GTGGGTTGATTGACGACCTACAGTGTACAGTCTGTATTAAGTTGTTAATAAAGTACA 1658
Db 77 GTGGGTTGATTGACGACCTACAGTGTACAGTCTGTATTAAGTTGTTAATAAAGTACA 18
Oy 1659 TGTAAACTTAAAAAA 1675
Db 17 TGTAAACTTAAAAAA 1

RESULT 16
AI249788/c 524 bp mRNA EST 03-FEB-1999
LOCUS qx51c10.x1 NCI-CGAP_Pan1 Homo sapiens cDNA clone IMAGE:2004862 3'
DEFINITION similar to gb:L06797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCRI
HOMOLOG (HUMAN);, mRNA sequence.
ACCESSION AI249788
NID g3846317
VERSION AI249788.1 GI:3846317
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 524)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLES National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Aug 21, 1998 this sequence version replaced.

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Unknown library type
Insert Length: 1614 Std Error: 0.00
Seq primer: -400P from Gibco
High quality sequence stop: 414.
Location/Qualifiers
1. 524
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2004862"
/clone_lib="NCI-CGAP_Pan1"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/notes="Organ: pancreas; Vector: pCMV-Sport6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.72 kb. Life Technologies catalog #:
11548-013"

BASE COUNT 195 a 103 c 73 g 151 t 2 others
```

```
ORIGIN
Query Match 25 0%; Score 435; DB 44; Length 524;
Best Local Similarity 100.0%; Pred. No. 2.1e-178;
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1246 TATGCTGTGGATTTTGTCTGTGTTCTTAACTTTTGTGAAGTTAATGACTTA 1305
Db 435 TATGCTGTGGATTTTGTCTGTGTTCTTAACTTTTGTGAAGTTAATGACTTA 376
Oy 1306 TTTATATTAATTTTTTTTGTTCATATTTGATGTGTCTAGGAGACCTGTGCGCAAGT 1365
Db 375 TTTATATTAATTTTTTTTGTTCATATTTGATGTGTCTAGGAGACCTGTGCGCAAGT 316
Oy 1366 TCTTAGTGTGTATGTCTGTGTAGAGCTGAGAAAAGGAACTGAACATTCCAGAGC 1425
Db 315 TCTTAGTGTGTATGTCTGTGTAGAGCTGAGAAAAGGAACTGAACATTCCAGAGC 256
Oy 1426 GTGTAGTAATCACTAAGCTAAGTAATGATCCCGAGCTGTTATGCATATATCTCT 1485
Db 255 GTGTAGTAATCACTAAGCTAAGTAATGATCCCGAGCTGTTATGCATATATCTCT 196
Oy 1486 CCAATCCCGTGAACGTTTTTCTGTCTTAAGACGTGATTTTGTCTGTAAGAGTGGCAC 1545
Db 195 CCAATCCCGTGAACGTTTTTCTGTCTTAAGACGTGATTTTGTCTGTAAGAGTGGCAC 136
Oy 1546 TTTAATCCAAAGCCCAAGTGTATAGAAATGCTGTTTTCAGTTTTCAGAGTGGGTT 1605
Db 135 TTTAATCCAAAGCCCAAGTGTATAGAAATGCTGTTTTCAGTTTTCAGAGTGGGTT 76
Oy 1606 GATTTCAGACCTACAGTGTACAGTCTGTATTAAGTTGTTAATAAAGTACTGTTAA 1665
Db 75 GATTTCAGACCTACAGTGTACAGTCTGTATTAAGTTGTTAATAAAGTACTGTTAA 16
Oy 1666 CTTAAAAA 1680
Db 15 CTTAAAAA 1

RESULT 17
AA259199/c 428 bp mRNA EST 20-AUG-1997
LOCUS nc17B08.r1 NCI-CGAP_Pri1 Homo sapiens cDNA clone IMAGE:1008375
DEFINITION similar to gb:L06797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCRI
HOMOLOG (HUMAN);, mRNA sequence.
ACCESSION AA259199
NID g1894641
VERSION AA259199.1 GI:1894641
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 428)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLES National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1394051.

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuquai,
M.D., Michael Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: David B. Kitzman, Ph.D.
CDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
```


QY 1546 TTATACCAAGCCCAAGTGTATAGAAATGCTGTTTTCAGTTTTCAGAGTGGTT 1605
|||||
Db 127 TTATACCAAGCCCAAGTGTATAGAAATGCTGTTTTCAGTTTTCAGAGTGGTT 68
QY 1606 GATTTCAGCAGCTGACGTGACGTGTTGTTTAAAGTAAAGTACAGTTAA 1665
|||||
Db 67 GATTTCAGCAGCTGACGTGACGTGTTGTTTAAAGTAAAGTACAGTTAA 8
QY 1666 CTTAAAA 1672
|||||
Db 7 CTTAAAA 1
RESULT 19
AI022030/c 535 bp mRNA EST 28-AUG-1998
LOCUS OW72a11.x1 Soares_fetal_liver_spleen_INFLS.S1 Homo sapiens CDNA
DEFINITION c72a11.x1 Soares_fetal_liver_spleen_INFLS.S1 Homo sapiens CDNA
PROTEIN-COUPLED RECEPTOR LCRI HOMOLOG (HUMAN); mRNA sequence.
ACCESSION AI022030
NID 93239383
VERSION AI022030.1 GI:3239383
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 535)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Jan 9, 1998 this sequence version replaced gi:785673.
CONTACT: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 975 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 189.
Location/Qualifiers
1..535
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="IMAGE:1652348"
/clone_id="Soares_fetal_liver_spleen_INFLS.S1"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: Liver and Spleen; Vector: pT73D (Pharmacia)
with a modified polylinker; Site.1: Pac I; Site.2: Eco RI;
This is a subtracted version of the original Soares fetal
liver spleen INFLS library. 1st strand cDNA was primed
with a Pac I - oligo(dT) primer [5',
AAGTGAAGATTAATTAAGATCTTTTATTTTATTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 205 a 108 c 79 g 142 t 1 others
ORIGIN

Query Match 24.4%; Score 423; DB 41; Length 535;
Best Local Similarity 99.8%; Pred. No. 3.3e-173;
Matches 473; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1195 TAAGTTACACATTTTTCAGATATAAAGACGACCAATATTGTACAGTTTATTCCTTG 1254
|||||
Db 474 TAAGTTACACATTTTTCAGATATAAAGACGACCAATATTGTACAGTTTATTCCTTG 415

QY 1255 TTGATTTTTCGTCGTGCTTCTTTAGTTTGTGGAAGTTAATGACTATTATTAATA 1314
|||||
Db 414 TTGGATTTTTCGTCGTGCTTCTTTAGTTTGTGGAAGTTAATGACTATTATTAATA 355
QY 1315 ATTTTTCGTTTCATATTGATGTGTCTAGCAGGACCTGTGGCCAGTTCTTACTTG 1374
|||||
Db 354 ATTTTTCGTTTCATATTGATGTGTCTAGCAGGACCTGTGGCCAGTTCTTACTTG 295
QY 1375 CTGATGTCTCTGCTGTAGGACTGTAGAAAGGAGACATTCGACAGCGTAGTGA 1434
|||||
Db 294 CTGATGTCTCTGCTGTAGGACTGTAGAAAGGAGACATTCGACAGCGTAGTGA 235
QY 1435 ATCAGCTAAAGCTAGATATATATCCCGAGCTGTTATGATATATCTCCATTCCG 1494
|||||
Db 234 ATCAGCTAAAGCTAGATATATATCCCGAGCTGTTATGATATATCTCCATTCCG 175
QY 1495 TGAAGCTTTTCTGCTGTTCTTAAGACGTGATTTGCTGTAGAGATGCGACTTAAACA 1554
|||||
Db 174 TGAAGCTTTTCTGCTGTTCTTAAGACGTGATTTGCTGTAGAGATGCGACTTAAACA 115
QY 1555 AAGCCCAAGCTGATAGAAATGCTGTTTTCAGTTTTCAGAGTGGTGTGATTGACG 1614
|||||
Db 114 AAGCCCAAGCTGATAGAAATGCTGTTTTCAGTTTTCAGAGTGGTGTGATTGACG 55
QY 1615 ACCTACAGTGTACAGCTCTGTTATTAAGTTGTTAATAAAGTACATGTTAACTT 1668
Db 54 ACCTACAGTGTACAGCTCTGTTATTAAGTTGTTAATAAAGTACATGTTAACTT 1
RESULT 20
AA148292/c 583 bp mRNA EST 06-AUG-1997
LOCUS z045e12.s1 Strathgane endothelial cell 937223 Homo sapiens CDNA
DEFINITION clone IMAGE:589870.3' similar to gb:L06797 PROBABLE G
PROTEIN-COUPLED RECEPTOR LCRI HOMOLOG (HUMAN); mRNA sequence.
ACCESSION AA148292
NID 91717715
VERSION AA148292.1 GI:1717715
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 583)
Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,
Chissee, S., Dietrich, N., Dubuque, T., Favell, A., Gish, W.,
Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N.,
Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,
Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Tinkley, M., J.,
Trevaskis, E., Underwood, K., Wohldmann, P., Waterston, R., Wilson, R.
and Marra, M.
Generation and analysis of 280,000 human expressed sequence tags
JOURNAL Genome Res. 6 (9), 807-828 (1996)
MEDLINE 97044478
COMMENT On Sep 12, 1996 this sequence version replaced gi:1394539.
CONTACT: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 888 Std Error: 0.00
Seq primer: -40m13 fwd. from Amersham
High quality sequence stop: 386.
Location/Qualifiers
1..583
/organism="Homo sapiens"
/db_xref="GDB:4621256"
/db_xref="taxon:9606"

```
/clone_lib="Stratagene endothelial cell 937223"  
/dev_stage="umbilical vein, 1 passage"  
/lab_host="50LR (kanamycin resistant)"  
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:  
XhoI; Cloned unidirectionally. Primer: Oligo dt.  
Umbilical vein endothelial cells, passed once. Average  
insert size: 1.0 kb; Uni-ZAP XR Vector: -5' adaptor  
sequence: 5' GAATTCGGCAGCAG 3' -3' adaptor sequence: 5'  
CTCGAGCTTTTCTTTTCTTTT 3'"
```

```
BASE COUNT      212 a      119 c      90 g      160 t      2 others  
ORIGIN
```

Query Match 24.2%: Score 421; DB 29; Length 583;

Best Local Similarity 100.0%; Pred. No. 2.4e-172;
Matches 421; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 1251 CTGTGATGATTTTGTCTGTTCTTTAGTTTGTGTAAGTTAATTGACTATTAT 1310  
|||||  
DB 421 CTGTGATGATTTTGTCTGTTCTTTAGTTTGTGTAAGTTAATTGACTATTAT 362  
OY 1311 ATAAATTTTTTGTTCATATTGATGTGTCTAGGACAGCTGTGGCCAAAGTTCTTA 1370  
|||||  
DB 361 ATAAATTTTTTGTTCATATTGATGTGTCTAGGACAGCTGTGGCCAAAGTTCTTA 302  
OY 1371 GTTGCTGATGCTGCTGTGTAGAGAGCTGAAGGAAGCACTTCCAGAGCGTGTA 1430  
|||||  
DB 301 GTTGCTGATGCTGCTGTGTAGAGAGCTGAAGGAAGCACTTCCAGAGCGTGTA 242  
OY 1431 GTGAATCAGGTAAGCTAGAAATGATCCCGAGCTGTTATGCATAGATATCTCCATT 1490  
|||||  
DB 241 GTGAATCAGGTAAGCTAGAAATGATCCCGAGCTGTTATGCATAGATATCTCCATT 182  
OY 1491 CCCGTGGAACGTTTTTCCGTCTCTTAAGACGTATTTGCTGTAAGAGATGCGACTATA 1550  
|||||  
DB 181 CCCGTGGAACGTTTTTCCGTCTCTTAAGACGTATTTGCTGTAAGAGATGCGACTATA 122  
OY 1551 ACCAAAGCCCAAGTGTATAGAAATGCTGTTTCAGTTTCAGAGAGGGTGATT 1610  
|||||  
DB 121 ACCAAAGCCCAAGTGTATAGAAATGCTGTTTCAGTTTCAGAGAGGGTGATT 62  
OY 1611 CAGCAGCTACAGTGTACAGTCTTTGATTAAGTTGTTAAATAAAGTACATGTTAACTTAA 1670  
|||||  
DB 61 CAGCAGCTACAGTGTACAGTCTTTGATTAAGTTGTTAAATAAAGTACATGTTAACTTAA 2  
OY 1671 A 1671  
DB 1 A 1
```

```
RESULT 21  
A1284206/c 419 bp mRNA EST 27-JAN-1999  
LOCUS q12b07.x1 Soares_NhmPpu_S1 Homo sapiens cDNA clone IMAGE:1857661  
DEFINITION 3' similar to gb:U06797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCRI  
HOMOLOG (HUMAN);, mRNA sequence.  
ACCESSION A1284206  
NID q3922439  
VERSION A1284206.1 GI:3922439  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
Eutheria; Primates; Catarrhini; Homiidae; Homo.
```

```
REFERENCE 1 (bases 1 to 419)  
AUTHORS NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT On Jan 14, 1998 this sequence version replaced gi:1798193.
```

Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Insert Length: 509 Std Error: 0.00
Seq primer: -40UP from Gibco.

FEATURES
Source Location/Qualifiers
1..419

```
/organism="Homo sapiens"  
/db_xref="taxon:9606"  
/map="16"  
/clone_lib="IMAGE:1857661"  
/clone_lib="Soares_NhmPpu_S1"  
/tissue_type="pooled human melanocyte, fetal heart, and  
pregnant uterus"  
/note="Organ: mixed (see below); Vector: pT73D-Pac  
(Pharmacia) with a modified polylinker; Site_1: Not I;  
Site_2: Eco RI; Equal amounts of plasmid DNA from three  
normalized libraries (melanocyte 2NbHM, pregnant uterus  
NBHPU, and fetal heart NBH19W) were mixed, and ss circles  
were made in vitro. Following HAP purification, this DNA  
was used as tracer in a subtractive hybridization  
reaction. The driver was PCR-amplified cDNAs from pools of  
5,000 clones made from the same 3 libraries. The pools  
consisted of I.M.A.G.E. clones 260232-265223,  
340488-345479, and 484488-489479."
```

```
BASE COUNT      158 a      92 c      61 g      108 t  
ORIGIN
```

Query Match 24.1%: Score 419; DB 44; Length 419;

Best Local Similarity 100.0%; Pred. No. 1.9e-171;
Matches 419; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 1252 TTGTTGATTTTGTCTGTTCTTTAGTTTGTGAGTTTAATGACTATTATTA 1311  
|||||  
DB 419 TTGTTGATTTTGTCTGTTCTTTAGTTTGTGAGTTTAATGACTATTATTA 360  
OY 1312 TAAATTTTTTTTTCATATTGATGTGTCTAGCAGAGCCTGTGGCCAAAGTTCTTAG 1371  
|||||  
DB 359 TAAATTTTTTTTTCATATTGATGTGTCTAGCAGAGCCTGTGGCCAAAGTTCTTAG 300  
OY 1372 TTGCTGTATGTCGCTGTGTAGAGACTGTAGAAAAGGAACATTCAGAGCGTGTAG 1431  
|||||  
DB 299 TTGCTGTATGTCGCTGTGTAGAGACTGTAGAAAAGGAACATTCAGAGCGTGTAG 240  
OY 1432 TGAATCAGCTAAAGCTAGAAATGATCCCGAGCTGTTATGATAGATTAATCTTCATT 1491  
|||||  
DB 239 TGAATCAGCTAAAGCTAGAAATGATCCCGAGCTGTTATGATAGATTAATCTTCATT 180  
OY 1492 CCGTGAAGGTTTTTCTGTTCTTAAGAGCTGATTTTGTCTGTAAGAGATGGCACTTATA 1551  
|||||  
DB 179 CCGTGAAGGTTTTTCTGTTCTTAAGAGCTGATTTTGTCTGTAAGAGATGGCACTTATA 120  
OY 1552 CCAAGCCCAAGGTGTATAGAAATCTGCTGTTTTCAGTTTCAGAGAGTGGTTGATTTC 1611  
|||||  
DB 119 CCAAGCCCAAGGTGTATAGAAATCTGCTGTTTTCAGTTTCAGAGAGTGGTTGATTTC 60  
OY 1612 AGCAGCTACAGTGTACAGCTTGTATTAAAGTTGTTAAATAAAGTACATGTTAACTTAA 1670  
|||||  
DB 59 AGCAGCTACAGTGTACAGCTTGTATTAAAGTTGTTAAATAAAGTACATGTTAACTTAA 1
```

```
RESULT 22  
AA490436 409 bp mRNA EST 15-AUG-1997  
LOCUS aa490436.1 NCI-CCAP_GCB1 Homo sapiens cDNA clone IMAGE:824478 5'  
DEFINITION similar to gb:U06797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCRI  
HOMOLOG (HUMAN);, mRNA sequence.  
ACCESSION AA490436  
NID g2219609  
VERSION AA490436.1 GI:2219609
```

KEYWORDS	EST
SOURCE	human
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 (bases 1 to 409)
TITLE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .
JOURNAL	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
COMMENT	Unpublished (1997) On Sep 12, 1996 this sequence version replaced gi:1393094.
FEATURES	<p>Contact: Robert Strausberg, Ph.D. Tel: (301) 496-1550 Email: Robert.Strausberg@nih.gov Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman, Ph.D., Gerald Marti, M.D. cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D. cDNA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: www.bio.lnl.gov/dbtp/image/image.html</p> <p>Seq primer: -28ml3 rev1 ET from Amersham.</p> <p>Location/Qualifiers</p> <p>1..409 /organism="Homo sapiens" /db_xref="taxon:9606" /map="20" /clone="IMAGE:824478" /clone.lib="NCI-CGAP_GCB1" /tissue_type="germinal center B cell" /lab_host="DH10B" /note="Vector: pT7T3-Pac (Pharmacia) with a modified polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from human tonsillar cells enriched for germinal center B cells by flow sorting (CD20+, IgD-), provided by Dr. Louis M. Staudt (NCI). Dr. David Allman (NCI) and Dr. Gerald Marti (CBBR). cDNA synthesis was primed with a Not I - oligo(dT) primer 15'-TGTTACCAATCTGAAGTGGAGGCGGCCGCTATTTTCTTTT-3' (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."</p>
BASE COUNT	102 a 74 c 77 g 156 t
ORIGIN	
Query Match	23.5%: Score 409; DB 34; Length 409;
Best Local Similarity	100.0%: Pred. No. 4e-167;
Matches 409; Conservative	0; Mismatches 0; Indels 0; Gaps 0.
993	CTATGCTTTCCTGGAGCCAAATTTAAACCTCTGCCAGCAGCAGCTACCTGTGAG 1052
1	CTATGCTTTCCTGGAGCCAAATTTAAACCTCTGCCAGCAGCAGCTACCTGTGAG 60
1053	CAGAGGCTCAGGCTCAGATCTCTCCAAAGGAAAGCGAGGTGACATTCGTGTTTC 1112
61	CAGAGGCTCAGGCTCAGATCTCTCCAAAGGAAAGCGAGGTGACATTCATCTGTTTC 120
1113	CAGTGAAGTCTGAGTCTTCAAGTTTCACTCAGCTAACACAGATGTTAAAGACTTTTTT 1172
121	CAGTGAAGTCTGAGTCTTCAAGTTTCACTCAGCTAACACAGATGTTAAAGACTTTTTT 180
1173	TATAGCATTAATTAAGTTTTTTTAAAGTACACATTTTTCAGATATATAAGACTGACCAAT 1232
181	TATAGCATTAATTAAGTTTTTTTAAAGTACACATTTTTCAGATATATAAGACTGACCAAT 240
1233	ATTGACAGTTTTTATTCGCTTGTGGAATTTTGTCTGTGTTCTTTAGTTTTTTGTGAAG 1292

Db	241	ATTGTACAGTTTATTTGCTTGTTGGAGATTTTTCCTTGCTGTTCTTAGTTTTTGGAAAG	300
OY	1293	TTTAATGACTTATTTAATAATTTTGTTCATATTGATGTCGTCTAGGCAGCA	1352
Db	301	TTTAATTACTTATTTATATAATTTTTTGTTCATATTGATGTCGTCTAGGCAGCA	360
OY	1353	CCTGTGGCCAAGTCTTAGTTCGTGAATGTCGTCTGCTAGCAGCTGTACA	1401
Db	361	CCTGTGGCCAAGTCTTAGTTCGTGAATGTCGTCTGCTAGCAGCTGTAGA	409
RESULT	23		
LOCUS	A1370816/c		
DEFINITION	A1370816	465 bp mRNA EST	16-FEB-1999
ACCSSION	U0149569		
NID	94149569		
VERSION	A1370816.1	GI:4149569	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;		
AUTHORS	Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap. National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index Unpublished (1997)		
JOURNAL	On Jan 17, 1998 this sequence version replaced gi:2043744.		
COMMENT	Contact: Robert Strausberg, Ph.D. Tel: (301) 496-1550 Email: Robert_Strausberg@nih.gov This clone is available royalty-free through LNL ; contact the IMAGE Consortium (infoimage.llnl.gov) for further information. Insert length: 913 Std Error: 0.00 Seq primer: -40UP from Gibco High quality sequence stop: 363. Location/Qualifiers 1..465 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:2048274" /clone_lib="Soares-total_fetus_ND2HF8_9w" /dev_stage="8-9 weeks" /lab_host="DH10B" /note="Vector: pRT3D-Pac (Pharmacia) with a modified polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from pooled 8-9 week (total) fetus material with a Not I - oligo(dT) primer [5' TGTTACCATCTGAAGGCGGCCGCGCTAATTTTTTTTTTTTTTTT 3'] . Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pRT3D vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."		
BASE COUNT	169 a	99 c	69 g 128 t
ORIGIN			
Query Match	Best Local Similarity	23.5%; Score 409; DB 45; Length 465; 100.0%; Pred. No. 3.9e-167;	
Matches	409; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
OY	1265	GCTCTGTGTTCTTCTTAGTTCGTGTAAGCTTATGACTTATTAATAATTTTTTTG	1324
Db	409	GCTCTGTGTTCTTCTTAGTTCGTGTAAGCTTATGACTTATTAATAATTTTTTTG	350
OY	1325	TTTCATATGATGATGTGCTAGACAGACCTGTGGCAAGTCTAGTGTGATGTCT	1384

Db 349 TTATATATGATGTCGTCTGAGCAGACCTGTGGCCAAAGTTCTTAGTCTGATATCT 290

QY 1385 CGTGTAGAGACTGTAGAAAAAGGAAGTGAACATTCACAGAGCGGTAGTGAATCAGCTAAA 1444

Db 289 CGTGTAGAGACTGTAGAAAAAGGAAGTGAACATTCACAGAGCGGTAGTGAATCAGCTAAA 230

QY 1445 GCTAGAAATGATCCCGAGCTTTTATGATAGATATATCTCTCCATCTCCCGTGGAAAGCTTT 1504

Db 229 GCTAGAAATGATCCCGAGCTTTTATGATAGATATATCTCTCCATCTCCCGTGGAAAGCTTT 170

QY 1505 TTCTGTCTTCTTAAGACGATTTTGTCTGTAGAGATGACACTTATATACCAAGCCCAAG 1564

Db 159 TTCTGTCTTCTTAAGACGATTTTGTCTGTAGAGATGACACTTATATACCAAGCCCAAG 110

QY 1565 TGTATAGAAATGCTGCTTTTTCAGTTTTCAGAGTGGGTGATTTTCAGACCTACAGCTG 1624

Db 109 TGTATAGAAATGCTGCTTTTTCAGTTTTCAGAGTGGGTGATTTTCAGACCTACAGCTG 50

QY 1625 TACGATCTGTATTAAGTGTATATAAAGTACATGTAACTTAAAAA 1673

Db 49 TACGATCTGTATTAAGTGTATATAAAGTACATGTAACTTAAAAA 1

RESULT 24
AA760743/c
LOCUS
DEFINITION
AA760743 509 bp mRNA EST 18-FEB-1998
n210809.s1 NCI-CGAP_GCB1 Homo sapiens CDNA clone IMAGE:1287353 3'
similar to gb:106797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCR1
HOMOLOG (HUMAN): mRNA sequence.

ACCESSION
NID
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
Ph.D., Gerald Marti, M.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldio, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bdrip/image/image.html

Insert length: 739 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 340.
Location/Qualifiers
1. 509
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1287353"
/clone_1b="NCI-CGAP_GCB1"
/tissue_type="germinal center B cell"
/lab_host="DH10B"
/note="Vector: p713D-Pac (Pharmacia) with a modified
polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, IgD+),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman

(NCI) and Dr. Gerald Marti (CEBR). CDNA synthesis was
primed with a Not I - 0.190(dt) primer
[5'-TGTACCAATCTGAGTGGAGCCGCTCATTTTCTTTTCTTTT-
3']. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified p713 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldio."

BASE COUNT 188 a 99 c 73 g 149 t
ORIGIN

Query Match 23.28; Score 403; DB 38; Length 509;
Best Local Similarity 100.0%; Pred. No. 1.5e-164;
Matches 403; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1279 TAGTTTGTGAAGTTTATGACTTATTTATATAATTTTGTTCATATGATGT 1338

Db 403 TAGTTTGTGAAGTTTATGACTTATTTATATAATTTTGTTCATATGATGT 344

QY 1339 GTGCTAGCAGACCTGTGGCCAAAGTTCTTAGTCTGTAATGCTGTAGACCTGT 1388

Db 343 GTGCTAGCAGACCTGTGGCCAAAGTTCTTAGTCTGTAATGCTGTAGACCTGT 284

QY 1399 AGAAAGGAGACTGAACATTCACAGCGCTAGTGAATCAGCTAAGTAAATGATCC 1458

Db 283 AGAAAGGAGACTGAACATTCACAGCGCTAGTGAATCAGCTAAGTAAATGATCC 224

QY 1459 CCAGCTGTTTATGATAGATATCTCTCCATCTCCGTTGAGACGTTTCTGTTCTTAAG 1518

Db 223 CCAGCTGTTTATGATAGATATCTCTCCATCTCCGTTGAGACGTTTCTGTTCTTAAG 164

QY 1519 ACGTATTTTCTGTAGAAATGCGACTTATACCAAGCCCAAGTGTATAGAAATGC 1578

Db 163 ACGTATTTTCTGTAGAAATGCGACTTATACCAAGCCCAAGTGTATAGAAATGC 104

QY 1579 TGGTTTTCAGTTTTCAGAGTGGGTGATTTTCAGACCTACAGTGTGATTT 1638

Db 103 TGGTTTTCAGTTTTCAGAGTGGGTGATTTTCAGACCTACAGTGTGATTT 44

QY 1639 AAGTGTATATAAAGTACATGTATAAATTAATAAATAAATAA 1681

Db 43 AAGTGTATATAAAGTACATGTATAAATTAATAAATAAATAA 1

RESULT 25
A1358446/c
LOCUS
DEFINITION
A1358446 417 bp mRNA EST 15-FEB-1999
qx19c11.x1 NCI CGAP Lym12 Homo sapiens CDNA clone IMAGE:2001812 3'
similar to gb:106797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCR1
HOMOLOG (HUMAN): mRNA sequence.

ACCESSION
NID
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Life Technologies catalog #: 11547-015
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:

www.bio.llnl.gov/db/p/image/image.html

Insert Length: 899 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 321.
Location/Qualifiers

FEATURES

Source

1. 417
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2001812"
/tissue_type="Lymphoma, follicular mixed small and large cell"
/lab_host="DH10B"
/note="Organ: lymph node; Vector: pCMV-SPORT6; Site: 1: SalI; Site: 2: NotI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.25 kb. Life Technologies catalog #: 11547-015"

BASE COUNT 148 a 88 c 61 g 120 t
ORIGIN

Query Match 23.2%; Score 403; DB 45; Length 417;
Best Local Similarity 100.0%; Pred. No. 1.6e-164;
Matches 403; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1280 AGTTTGTGAAGTTAATGACTATTATATAATTTTGTTCATATGATGTC 1339
DB 403 AGTTTGTGAAGTTAATGACTATTATATAATTTTGTTCATATGATGTC 344

QY 1340 TGTCTAGGAGACCGTGGCCAGTTCTAGTGTGCTGATGCTGCTGATGACTGTA 1399
DB 343 TGTCTAGGAGACCGTGGCCAGTTCTAGTGTGCTGATGCTGCTGATGACTGTA 294

QY 1400 GAAAGGAGACTGACATTCACAGCGGTAGTGAATCAGTAAGCTAGAAATGATCC 1459
DB 283 GAAAGGAGACTGACATTCACAGCGGTAGTGAATCAGTAAGCTAGAAATGATCC 224

QY 1460 CAGCTGTTATGATGATATATCTCTCCATCCCGTGAACGTTTCCGTCTTAAGA 1519
DB 223 CAGCTGTTATGATGATATATCTCTCCATCCCGTGAACGTTTCCGTCTTAAGA 164

QY 1520 CGGATTTGCTGTAGAAGTGCACCTATATACCAAGTGTATGAAGATGCT 1579
DB 163 CGGATTTGCTGTAGAAGTGCACCTATATACCAAGTGTATGAAGATGCT 104

QY 1580 GGTTTTCAGTTTTCAGAGTGGGTGATTTTCAGACCTACAGTGTACGCTTGATTA 1639
DB 103 GGTTTTCAGTTTTCAGAGTGGGTGATTTTCAGACCTACAGTGTACGCTTGATTA 44

QY 1640 AGTTGTTAATAAAGTACATGTTAACTTAAAAA 1682
DB 43 AGTTGTTAATAAAGTACATGTTAACTTAAAAA 1

RESULT 26
AI660355 401 bp mRNA EST 10-MAY-1999
LOCUS
DEFINITION we62b07.x1 Soares_thymus_NHFTb Homo sapiens cDNA clone
IMAGE:2345653 3' similar to gb:106797 PROBABLE G PROTEIN-COUPLED
RECEPTOR LCRI HOMOLOG (HUMAN); mRNA sequence.
ACCESSION AI660355
NID 94763925
VERSION AI660355.1 GI:4763925
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 401)
NCI-CCG http://www.ncbi.nlm.nih.gov/ncicgaf.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index

JOURNAL COMMENT

Unpublished (1997)
On Mar 10, 1998 this sequence version replaced gi:2949344.

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40UP from Gibco.

FEATURES

source

1. 401
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2345653"
/dev_stage="fetal" thymus_NHFTb"
/lab_host="DH10B (phage-resistant)"
/note="Organ: thymus; pooled; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site: 1: Not I; Site: 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dt) primer [5', TGTACCATCTGAGAGTGGAGCGCGCCAGCTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 107 a 59 c 82 g 153 t
ORIGIN

Query Match 23.1%; Score 401; DB 49; Length 401;
Best Local Similarity 100.0%; Pred. No. 1.2e-163;
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1181 AATAACTTTTATAGTTACACATTTTTCAGATATATAAGACGTACCAATATGTACA 1240
DB 1 AATAACTTTTATAGTTACACATTTTTCAGATATATAAGACGTACCAATATGTACA 60

QY 1241 GTTTTATGCTGTGGATTTTGTCTGTGTTCTTATGTTTGTGAAGTTAATG 1300
DB 61 GTTTTATGCTGTGGATTTTGTCTGTGTTCTTATGTTTGTGAAGTTAATG 120

QY 1301 ACTTATTATTAATTTTGTTCATATGATGCTGTCTAGCAGACCTGTGCG 1360
DB 121 ACTTATTATTAATTTTGTTCATATGATGCTGTCTAGCAGACCTGTGCG 180

QY 1361 CAAGTCTTAGTCTGTATGCTGTGTAGACCTGTAGAAAAGGACCTGAACATTCC 1420
DB 181 CAAGTCTTAGTCTGTATGCTGTGTAGACCTGTAGAAAAGGACCTGAACATTCC 240

QY 1421 AGAGCGTAGTGAATCAGTAAAGCTAGAAATGATCCAGCTGTTATGCTATAGTA 1480
DB 241 AGAGCGTAGTGAATCAGTAAAGCTAGAAATGATCCAGCTGTTATGCTATAGTA 300

QY 1481 TCTTCATATCCCGTGAACGTTTCTGTTCTTAAGACGTGATTTTGTGTAAGAT 1540
DB 301 TCTTCATATCCCGTGAACGTTTCTGTTCTTAAGACGTGATTTTGTGTAAGAT 360

QY 1541 GGCACATTATACCAAGCCCAAGTGTATGAATGCTGG 1581
DB 361 GGCACATTATACCAAGCCCAAGTGTATGAATGCTGG 401

RESULT 27
AA386000/c 393 bp mRNA EST 21-APR-1997
LOCUS
DEFINITION EST9758 Pancreas tumor I Homo sapiens cDNA 3' end similar to
guanine nucleotide-binding protein HM89, mRNA sequence.
ACCESSION AA386000
NID 52038337
VERSION AA386000.1 GI:2038337
KEYWORDS EST.

SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Mammalia: Eutheria: Primates: Catarrhini: Homiidae: Homo.
AUTHORS Adams,M.D., Kertlavage,A.R., Fleischmann,R.D., Fulton,R.A., Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White,O., Sutton,G., Blake,J.A., Brandon,R.C., Maniatis,C., Clifton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodex,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr., Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M., Moreno-Palancas,R.F., McDonald,L.A., Nauyen,D.T., Pelligrino,S.M., Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R., Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y., Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J., Dimke,D., Feng,D.-F., Ferrite,A., Fischer,C., Hastings,G.A., He,M.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H., Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M., and Venter,J.C.
TITLE Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence
JOURNAL Nature 377 (6547 Suppl.), 3-174 (1995)
MEDLINE 95026280
COMMENT On Dec 30, 1996 this sequence version replaced gi:1530936.
Other ESTs: EST9739 T0C16975
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699055
Fax: 3018699423
Email: arkerlav@tigr.org
For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
Seq primer: M13-21.

FEATURES
SOURCE Location/Qualifiers
1..393
/organism="Homo sapiens"
/db_xref="ATCC (Inhost):190162"
/db_xref="taxon:9606"
/clone_lib="Pancreas tumor I"
/dev_stage="adult"
/note="Organ: pancreas; Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT 142 a 86 c 60 g 105 t
ORIGIN

Query Match 22.6%; Score 393; DB 33; Length 393;
Best Local Similarity 100.0%; Pred. No. 3.4e-160;
Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1276 CTTTGGTTTGTGAAGTTAATGACTATTATTAATTTTGTTCATATTGA 1335
|||||
393 CTTTGGTTTGTGAAGTTAATGACTATTATTAATTTTGTTCATATTGA 334
OY 1336 TGTGTCTAGGACAGACTGTGGCCAAAGTCTTAGTCTGTATGCTGTGGTAGAC 1395
|||||
DB 333 TGTGTCTAGGACAGACTGTGGCCAAAGTCTTAGTCTGTATGCTGTGGTAGAC 274
OY 1396 TGTAAAAAGGAAGTGAATCTCCAGAGCGTGTAGTAATCAGCTAAGCTGAATGA 1455
|||||
DB 273 TGTAAAAAGGAAGTGAATCTCCAGAGCGTGTAGTAATCAGCTAAGCTGAATGA 214
OY 1456 TCCGAGCTGTATGATAGATTAATCTCCATTCCTGGAGAGCTTTTCTGTCTT 1515
|||||
DB 213 TCCGAGCTGTATGATAGATTAATCTCCATTCCTGGAGAGCTTTTCTGTCTT 154
OY 1516 AAGACGTGATTTTGTCTGTGAAGATGGCACTTAATACCAAGCCCAAGCTGTATGAAA 1575

|||||
DB 153 AAGACGTGATTTTGTCTGTAGAACATGGCACTTAATACCAAGCCCAAGGATATGAAA 94
OY 1576 TCGTGGTTTTCAGTTTTCAGGAGTGGGTGATTTTCAGCACCCTACAGTGTACAGTCTTGT 1635
|||||
DB 93 TCGTGGTTTTCAGTTTTCAGGAGTGGGTGATTTTCAGCACCCTACAGTGTACAGTCTTGT 34
OY 1636 ATTAAGTGTATTAATAAGTACATGTTAACTT 1668
|||||
DB 33 ATTAAGTGTATTAATAAGTACATGTTAACTT 1

RESULT 28
AA479467 521 bp mRNA EST 08-AUG-1997
LOCUS zvl7110.r1 Soares_NHMPu_S1 Homo sapiens cDNA clone IMAGE:753931 5'
DEFINITION similar to gb:106797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCRI
HOMOLOG (HUMAN);, mRNA sequence.
ACCESSION AA479467
NID 92208023
VERSION AA479467.1 GI:2208023
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Mammalia: Eutheria: Primates: Catarrhini: Homiidae: Homo.
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisler,G., Jost,S., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
TITLE WashU-Merck EST Project 1997
JOURNAL Unpublished (1997)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1393144.

CONTACT: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28m3 rev2 ET from Andersham
High quality sequence stop: 421.

FEATURES
SOURCE Location/Qualifiers
1..521
/organism="Homo sapiens"
/db_xref="GDB:5976803"
/db_xref="taxon:9606"
/map="956G09:1"
/clone="IMAGE:753931"
/clone_lib="Soares_NHMPu_S1"
/tissue_type="Pooled human melanocyte, fetal heart, and pregnant uterus"
/lab_host="DH10B"
/note="Organ: mixed (see below); Vector: p773D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Equal amounts of 2bDM, pregnant uterus NBHP, and fetal heart NBHN19W were mixed, and ss circles were made in vitro. Following NRP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of 1.M.A.G.E. clones 260232-265223, 340488-345479, and 484488-489479."
BASE COUNT 132 a 104 c 103 g 182 t
ORIGIN

Query Match 22.2%; Score 386; DB 34; Length 521;
Best Local Similarity 99.6%; Pred. No. 3.4e-157;

```

Insert Length: 623      Std Error: 0.00
Seq primer: -40m13 fwd. ET from AmerSham
High quality sequence stop: 319.
Location/Qualifiers
    1..504
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="IMAGE:170363Z"
        /clone_1ib="NCI_CGAP_Brn23"
        /tissue_type="gliblastoma (pooled)"
        /lab_host="DH10B"
        /note="Organ: brain; Vector: pTR73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATTCAGAACTGGAGCGGCCGCATATCTTTTTTTTTTTTTTTTTTTT T 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pTVT3 vector. Library is normalized, and was constructed by Bento Scores and M.Fatima Bonaldo."

BASE COUNT       185 a         102 c          71 g         146 t

ORIGIN
Query Match:           22.2%; Score 386; DB 43; Length 504;
Best Local Similarity 100.0%; Pred. No. 3,4e-157;
Matches   386; Conservative     0; Mismatches     0; Indels     0; Gaps     0;

QY   1296 AATTGACTATTATTATAAATTTTTTGTTCCATGTGATGCTGTAGCAGACCCT 1355
Db   386 AATTGACTATTATTATAAATTTTTTGTTCCATGTGATGCTGTAGCAGACCCT 327
QY   1356 GTGGCCAAGTCTTAGTGTGCTGTATGTCCTGTGTAGSACTGTAAAAAGGAAGTAAC 1415
Db   326 GTGGCCAAGTCTTAGTGTGCTGTATGTCCTGTGTAGSAGACTGTAAAAGAAGTAAC 267
QY   1416 ATTCCAGAGCGTGTAGTGAATCACGTAAAGCTAGAACATGATCCCAGCTGTTATGCATA 1475
Db   266 ATTCACAGAGCGTGTAGTGAATCACGTAAAGCTAGAACATGATCCCAGCTGTTATGCATA 207
QY   1476 GATATATCTCTCCATTCCCGGTGGAAAGCTTTTCCTGTTCTTAAGACGTGATTTGCTGTAG 1535
Db   206 GATATATCTCTCATTCCTCGGTGGAAAGCTTTTCCTGTTCTTAAGACGTGATTTGCTGTAG 147
QY   1536 AAGATGGCACTTATAACCAAAGCCCCAAAGTGSATAGAAATCCGTGTTTCAGTATTCA 1595
Db   146 AAGATGGCACTTATAACCAAAGCCCCAAAGTGSTATAGAAATCCGTGTTTCAGTATTCA 87
QY   1596 GGAGTGGGTTGATTTACAGACCCTACAGCTGTACAGTCTGTATTAAGTTGTTAATAAAAGT 1655
Db   86 GGAGTGGGTTGATTTACAGACCCTACAGTGTACAGTCTGTATTAAGTTGTTAATAAAAGT 27
QY   1656 ACATGTTAACTTAAAAAAAAAAAAA 1681
Db   26 ACATGTTAACTTAAAAAAAAAAAAA 1

RESULT 30
LOCUS AI193140/c 475 bp mRNA EST 29-OCT-1998
DEFINITION qe#0f12.s1 Soares_fetal_lung_Nbh1l9w Homo sapiens CDNA clone IMAGE:1741487 3 similar to gp:L06797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCRI HOMOLOG (HUMAN); , mRNA sequence.
ACCESSION AI193140
NID 93744349
VERSION AI193140.1 GI:3744349
KEYWORDS EEST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Euheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 475)

```

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
JOURNAL Tumor Gene Index
COMMENT Unpublished (1997)
On Jan 14, 1998 this sequence version replaced gi:1797529.

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
This clone is available royalty-free through LNL: contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Insert Length: 911 Std Error: 0.00
Seq primer: -40m3 fwd. ET from AmerSham
High quality sequence stop: 329.
Location/Qualifiers

FEATURES
Source

1..475
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1741487"
/clone_lib="Soares_fetal_Lung_NbHL19w"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: Lung; Vector: pT73D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGAAGTGGAGCGCGCATTTTCTTTTCTTTT-3']
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M.Fatima Bonaldo. This library was constructed
from the same fetus as the fetal heart library, Soares
fetal heart NbHL19w."

BASE COUNT 176 a 101 c 73 g 125 t
ORIGIN

Query Match 22.2%; Score 386; DB 43; Length 475;
Best Local Similarity 100.0%; Pred. No. 3,5e-157;
Matches 386: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1284 TTGTGAGTTAATTGACTATTATATATAATTTTGTGTTTCATATTGATGTGTC 1343
|||||
DB 386 TTGTGAGTTAATTGACTATTATATATAATTTTGTGTTTCATATTGATGTGTC 327
OY 1344 TGGGAGGACCTGTGGCCAGTTTCTAGTGTCTGTCTGTGTGTAGGACGTGAAA 1403
|||||
DB 328 TAGGAGGACCTGTGGCCAGTTTCTAGTGTCTGTGTGTGTAGGACGTGAAA 267
OY 1404 AGGGAAGTGAACATTCAGAGCGTGTAGTAATCAGCTAAAGCTAGAACATCCCCAGC 1463
|||||
DB 266 AGGGAAGTGAACATTCAGAGCGTGTAGTAATCAGCTAAAGCTAGAACATCCCCAGC 207
OY 1464 TGTATTATGATGATATCTCTCCATCCCGTGGAAAGCTTTTCTGTTCTTAAGACGTG 1523
|||||
DB 206 TGTATTATGATGATATCTCTCCATCCCGTGGAAAGCTTTTCTGTTCTTAAGACGTG 147
OY 1524 ATTTGCTGTAGAGATGCTATTATACCAAGCCCAAGGTGTATAAATGCTGGTT 1583
|||||
DB 146 ATTTGCTGTAGAGATGCTATTATACCAAGCCCAAGGTGTATAAATGCTGGTT 87
OY 1584 TTTCAGTTTTCAGAGTGGTGTGATTTAGCAGCTACAGCTACAGCTTGTATTAGTT 1643
|||||
DB 86 TTTCAGTTTTCAGAGTGGTGTGATTTAGCAGCTACAGCTACAGCTTGTATTAGTT 27
OY 1644 GTTATAAAGTACATGTTAACTTA 1669
|||||
DB 26 GTTATAAAGTACATGTTAACTTA 1

RESULT 31
A1434652/C

LOCUS A1434652 531 bp mRNA EST 30-MAR-1999
DEFINITION t134d02.x1 NCI-CGAP_Lym12 Homo sapiens cDNA clone IMAGE:213235 3'
Similar to gb:D06797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCRI
HOMOLOG (HUMAN); mRNA sequence.

ACCESSION A1434652
NID 94297677
VERSION A1434652.1 GI:4297677
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 531)
NCT-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)

JOURNAL On May 7, 1998 this sequence version replaced gi:312121.
COMMENT

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Life Technologies catalog #: 11547-015
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www.bio.lnl.gov/dbfp/image/image.html

Insert Length: 742 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 293.
Location/Qualifiers

FEATURES
Source

1..531
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="12p"
/clone="IMAGE:2132355"
/clone_lib="NCI-CGAP_Lym12"
/tissue_type="lymphoma, follicular mixed small and large
cell"
/lab_host="DH10B"
/note="Organ: lymph node; Vector: pCMV-SPORT6; Site_1:
SalI; Site_2: NotI; Cloned unidirectionally. Primer:
oligo dT. Average insert size 1.25 kb. Life Technologies
catalog #: 11547-015"

BASE COUNT 194 a 105 c 73 g 158 t 1 others
ORIGIN

Query Match 21.6%; Score 375; DB 45; Length 531;
Best Local Similarity 100.0%; Pred. No. 2e-152;
Matches 375: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1316 TTTTGTTCATATTGATGTGTCTAGCAGACCTGTGCCAAGTTCTTAGTGC 1375
|||||
DB 375 TTTTGTTCATATTGATGTGTCTAGCAGACCTGTGCCAAGTTCTTAGTGC 316
OY 1376 TGTATGTCGTGGTGAAGCTAGAAAAGGAGACTGACATTCAGAGCGGTAGTGA 1435
|||||
DB 315 TGTATGTCGTGGTGAAGCTAGAAAAGGAGACTGACATTCAGAGCGGTAGTGA 256
OY 1436 TCACGTAAGCTAGAAATGATCCAGCTGTTATGATAGATTAATCTCCATTCCTG 1495
|||||
DB 255 TCACGTAAGCTAGAAATGATCCAGCTGTTATGATAGATTAATCTCCATTCCTG 196
OY 1496 GGAAGCTTTTCTGTTCTTAAGACGTGATTTTCTGTGAAGAATGGCACTTAACCA 1555
|||||
DB 195 GGAAGCTTTTCTGTTCTTAAGACGTGATTTTCTGTGAAGAATGGCACTTAACCA 136
OY 1556 AGCCCAAAAGTGTATAGAAATGCTGTTTTCAGTTTCAGAGTGGTGTGATTTTCAGA 1615
|||||
DB 135 AGCCCAAAAGTGTATAGAAATGCTGTTTTCAGTTTTCAGAGTGGTGTGATTTTCAGA 76

OY 1516 CCTACAGTGTACAGTCTTGTATTAGTTGTTAATAAAGTACATGTTAACTTAAAAA 1675
|||||
Db 75 CCTACAGTGTACAGTCTTGTATTAGTTGTTAATAAAGTACATGTTAACTTAAAAA 16
OY 1516 AAAAAAAAAAAAAA 1690
|||||
Db 15 AAAAAAAAAAAAAA 1
RESULT 32
AA824607 538 bp mRNA EST 18-MAR-1998
LOCUS OC83f08.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1356327
DEFINITION similar to gb:U06797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCRI
HOMOLOG (HUMAN);, mRNA sequence.
ACCESSION AA824607
NID 92896629
VERSION AA824607.1 GI:2896629
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 538)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Sep 29, 1997 this sequence version replaced gi:1520732.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
Ph.D., Gerald Marti, M.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www-bio.llnl.gov/bdrrp/image/image.html
Insert Length: 643 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 191.
Location/Qualifiers
1..538
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="1356327"
/clone_lib="NCI_CGAP_GCB1"
/tissue_type="germinal center B cell"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+ IgD-),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
primed with a Not I - oligo(dT) primer
15'-GTTACCAATCTGAAGTGGAGCGCCGCTTTTCTTTTCTTTT-
3'. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 185 a 102 c 85 g 166 t
ORIGIN
Query Match 21.1%; Score 366; DB 39; Length 538;
Best Local Similarity 100.0%; Pred. No. 1.5e-148;

Matches 366; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1325 TTTCATATTGANGTGTGTCTAGGACGACCTGGCCAGTCTTAGTGGTATGCT 1384
|||||
Db 366 TTTCATATTGANGTGTGTCTAGGACGACCTGGCCAGTCTTAGTGGTATGCT 307
OY 1385 CGTGGTAGACCTGTAGAAAAGGAAGTCAATCCAGAGCGGTAGTGAATCAGCTAAA 1444
|||||
Db 306 CGTGGTAGACCTGTAGAAAAGGAAGTCAATCCAGAGCGGTAGTGAATCAGCTAAA 247
OY 1445 GCTAGAAATGATCCCGACGCTGTTTATGATATATCTCTCATTCGCGTGAAGTTT 1504
|||||
Db 246 GCTAGAAATGATCCCGACGCTGTTTATGATATATCTCTCATTCGCGTGAAGTTT 187
OY 1505 TTCCTGTTCTTAAGACGATTTTGGCTGAGAAATGACGACTTAATACCAAGCCCAAG 1564
|||||
Db 186 TTCCTGTTCTTAAGACGATTTTGGCTGAGAAATGACGACTTAATACCAAGCCCAAG 127
OY 1565 TGGTATAGAAATGCTGTTTTCAGTTTTCAGAGAGTGGTGTGATTTTCAGCACCCTACAGTG 1624
|||||
Db 126 TGGTATAGAAATGCTGTTTTCAGTTTTCAGAGAGTGGTGTGATTTTCAGCACCCTACAGTG 67
OY 1625 TACAGTCTGTATTAGTGTATTAAAGTACATCTTAACCTTAAAAA 1684
|||||
Db 66 TACAGTCTGTATTAGTGTATTAAAGTACATCTTAACTTAAAAA 7
OY 1685 AAAAAA 1690
|||||
Db 6 AAAAAA 1
RESULT 33
AA479357/c
LOCUS AA479357
DEFINITION
AA479357 537 bp mRNA EST 08-AUG-1997
2117110.s1 Soares Nhlmpu.s1 Homo sapiens cDNA clone IMAGE:753931 3'
similar to gb:U06797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCRI
HOMOLOG (HUMAN);, mRNA sequence.
ACCESSION AA479357
NID 92207913
VERSION AA479357.1 GI:2207913
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 537)
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisler, G., Jost, S.,
Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J.,
Moore, B., Schellenberg, K., Stepien, M., Tan, F., Theising, B.,
White, Y., Wylie, T., Waterston, R. and Wilson, R.
WashU-Merck EST Project 1997
Unpublished (1997)
On Sep 12, 1996 this sequence version replaced gi:1393855.
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -4ml3 fwd. ET from Amersham
High quality sequence stop: 372.
Location/Qualifiers
1..537
/organism="Homo sapiens"
/db_xref="GDB:5976803"
/db_xref="taxon:9606"
/map="19p12-p13.1"
/clone_image="753931"
/clone_lib="Soares_Nhlmpu.s1"
/tissue_type="Pooled human melanocyte, fetal heart, and

pregnant uterus"
/lab_host="DH10B"
/note="Organ: mixed (see below); Vector: pT73D-Pac
(pharmacia) with a modified polylinker; Site_1: Not I;
Site_2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2NHM, pregnant uterus
NBHPV, and fetal heart NBH19M) were mixed, and ss circles
were made in vitro. Following HAP purification, this DNA
was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
5,000 clones made from the same 3 libraries. The pools
consisted of I.M.A.G.E. clones 260232-265223,
340488-345479, and 484488-489479."

BASE COUNT 194 a 107 c 89 g 147 t
ORIGIN

Query Match 20.4%; Score 354; DB 34; Length 537;
Best Local Similarity 100.0%; Pred. No. 2.4e-143;
Matches 354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1318 TTTTGTTCATATGATGTCGTCTAGCAGAGACCTGTGCCAGTTTATGTCCTG 1377
|||||
Db 354 TTTTGTTCATATGATGTCGTCTAGCAGAGACCTGTGCCAGTTTATGTCCTG 295
QY 1378 TATGTCGTGTAGACCTGTAGAAAAGGAGACTGAAACATTCAGAGCGGTGTGATC 1437
|||||
Db 294 TATGTCGTGTAGACCTGTAGAAAAGGAGACTGAAACATTCAGAGCGGTGTGATC 235
QY 1438 ACCTAAAGCTAGAAATGATCCCGAGCTGTTATGATAGATATCTCCATTCCTG 1497
|||||
Db 234 ACCTAAAGCTAGAAATGATCCCGAGCTGTTATGATAGATATCTCCATTCCTG 175
QY 1498 AACGTTTTCCTGTTCTTAAGAGCTATTTGCTGTAGAAAGTGGCACTTATACCAAG 1557
|||||
Db 174 AACGTTTTCCTGTTCTTAAGAGCTATTTGCTGTAGAAAGTGGCACTTATACCAAG 115
QY 1558 CCCAAGTGTGTAGAAATGCTGTTTTCAGTTTCAGAGTGGGTGTTTCAGCACC 1617
|||||
Db 114 CCCAAGTGTGTAGAAATGCTGTTTTCAGTTTCAGAGTGGGTGTTTCAGCACC 55
QY 1618 TACAGTGTACGCTTGTGTATTAAGTTGTTAATAAAGTACATGTTAACTTAA 1671
|||||
Db 54 TACAGTGTACGCTTGTGTATTAAGTTGTTAATAAAGTACATGTTAACTTAA 1

RESULT 34
A1682902/c 405 bp mRNA EST 26-MAY-1999
LOCUS A1682902
DEFINITION wc69g07.x1 NCI-CGAP_Pauli Homo sapiens cDNA clone IMAGE:2323932 3'
similar to gb:L06797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCR1
HOMOLOG (HUMAN); mRNA sequence.
ACCESSION A1682902
NID 94883084
VERSION A1682902.1 GI:4893084
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 405)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
TITLE Unpublished (1997)
JOURNAL On May 18, 1998 this sequence version replaced gi:3137611.
COMMENT

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Life Technologies catalog #: 11548-013
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/dbp/image/image.html

Seq primer: -40UP from Gibco
High quality sequence stop: 339.
Location/Qualifiers

FEATURES

Source
1..405
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="X"
/clone="IMAGE:2323932"
/clone_lib="NCI CGAP_Pauli"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/note="Organ: pancreas; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.72 kb. Life Technologies catalog #:
11548-013"

BASE COUNT 149 a 88 c 61 g 106 t 1 others
ORIGIN

Query Match 20.4%; Score 354; DB 50; Length 405;
Best Local Similarity 99.8%; Pred. No. 2.6e-143;
Matches 404; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1266 TCTTGTGTTCTTTAGTTTGTGAAATTAATGACTATTATATAATTTTGT 1325
|||||
Db 405 TCTTGTGTTCTTTAGTTTGTGAAATTAATGACTATTATATAATTTTGT 346
QY 1326 TTCAATATGATGTGTCTAGAGAGACCGTGGCCAGTCTTACTGCTATGCTC 1385
|||||
Db 345 TTCAATATGATGTGTCTAGAGAGACCGTGGCCAGTCTTACTGCTATGCTC 286
QY 1386 GTGTAGAGACTGTAGAAAAGGAGAACTGAACATTCAGAGCGGTGTGATCAGTAAAG 1445
|||||
Db 285 GTGTAGAGACTGTAGAAAAGGAGAACTGAACATTCAGAGCGGTGTGATCAGTAAAG 226
QY 1446 CTGAAATATGATCCCGAGCTGTTATGATAGATATCTCCATTCCTCGGAGACGTTT 1505
|||||
Db 225 CTGAAATATGATCCCGAGCTGTTATGATAGATATCTCCATTCCTCGGAGACGTTT 166
QY 1506 TCTGTCTTAAAGCTGATTTGCTGTAGAGATGGCACTTAAACCAAGCCCAAGT 1565
|||||
Db 165 TCTGTCTTAAAGCTGATTTGCTGTAGAGATGGCACTTAAACCAAGCCCAAGT 106
QY 1566 GGTATAGAAATGCTGTTTTCAGTTTTCAGAGTGGGTGATTTGAGCAGCTACAGTG 1625
|||||
Db 105 GGTATAGAAATGCTGTTTTCAGTTTTCAGAGTGGGTGATTTGAGCAGCTACAGTG 46
QY 1626 ACAGCTCTGTATTAAGTTGTTAATAAAGTACATGTTAACTTAA 1670
|||||
Db 45 ACAGCTCTGTATTAAGTTGTTAATAAAGTACATGTTAACTTAA 1

RESULT 35
AA707668/c 786 bp mRNA EST 24-DEC-1997
LOCUS AA707668
DEFINITION zj29f11.s1 Soares_fetal_liver.spleen.INFUS.S1 Homo sapiens cDNA
clone IMAGE:451727 3' similar to gb:L06797 PROBABLE G
PROTEIN-COUPLED RECEPTOR LCR1 HOMOLOG (HUMAN); mRNA sequence.
ACCESSION AA707668
NID 92717586
VERSION AA707668.1 GI:2717586
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 786)
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisler, G., Jost, S.,
Kizman, D., Kucada, T., Lacy, M., Le, N., Lennon, G., Marra, M.,
Martin, J., Moore, B., Schellenberg, K., Stepoe, M., Tan, F.,

TITLE
JOURNAL
COMMENT

Thising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
WashU-MCI human EST Project
Unpublished (1997)
On Nov 6, 1997 this sequence version replaced gi:315382.

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 268.

FEATURES

source
1. 786
/organism="Homo sapiens"
/db_xref="GDB:1388083"
/db_xref="taxon:9606"
/map="17:15q26.1"
/clone="IMAGE:451727"
/clone_lib="Soares_fetal_liver_spleen_INFLS_S1"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: Liver and Spleen; Vector: pT73D (Pharmacia)
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
This is a subcloned version of the original Soares fetal
liver spleen INFLS library. 1st strand cDNA was primed
with a Pac I - Oligo(dT) primer [5',
AACTGAGAGATTATTAAGATCTTTTCTTTTCTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT 256 a 156 c 163 g 211 t
ORIGIN

Query Match 20.0%; Score 347; DB 31; Length 786;
Best Local Similarity 99.6%; Pred. No. 2.3e-140;
Matches 447; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1220 AAGACTGACCAATTTGACAGTTTATTCCTGTTGGATTGTCCTGTTCTTT 1279
|||||
Db 449 AAGACTGACCAATTTGACAGTTTATTCCTGTTGGATTGTCCTGTTCTTT 390
QY 1280 AGTTTGTGAGTTTAACTGATTTATTAATAATTTTGTGTTCAATATGATGTG 1339
|||||
Db 389 AGTTTGTGAGTTTAACTGATTTATTAATAATTTTGTGTTCAATATGATGTG 330
QY 1340 TGTCTAGGACGACCTGTGGCCAACTTTAGTCTGTATGTCGCGGTGAGACTGTA 1399
|||||
Db 329 TGTCTAGGACGACCTGTGGCCAACTTTAGTCTGTATGTCGCGGTGAGACTGTA 270
QY 1400 GAAAAGGACATGACATTTCCAGAGCGTGTAGTATCAGCTAAGCTAGAATATGCC 1459
|||||
Db 269 GAAAAGGACATGACATTTCCAGAGCGTGTAGTATCAGCTAAGCTAGAATATGCC 210
QY 1460 CAGCTGTTATGATGATATCTCTCCATTCCTCCGTGGAAGCTTTTCTGTTCTTAAGA 1519
|||||
Db 209 CAGCTGTTATGATGATATCTCTCCATTCCTCCGTGGAAGCTTTTCTGTTCTTAAGA 150
QY 1520 CGTGAATTTGCTGTAGAGAGTGGCACTTATAACCAAGCCCAAGTGTATGAATGCT 1579
|||||
Db 149 CGTGAATTTGCTGTAGAGAGTGGCACTTATAACCAAGCCCAAGTGTATGAATGCT 90
QY 1580 GGTGTTTTCAGTTTTCAGAGGTGGTGTGATTCAGACCTACAGTGTACAGTCTTGATTA 1639
|||||
Db 89 GGTGTTTTCAGTTTTCAGAGGTGGTGTGATTCAGACCTACAGTGTACAGTCTTGATTA 30

QY 1640 AGTTGTTATATAAAGTACATGTTAACTT 1668
Db 29 AGTTGTTATATAAAGTACATGTTAACTT 1

RESULT 36
AA284857/c 454 bp mRNA EST 08-AUG-1997
LOCUS z122e05.s1 Soares ovary tumor NbhOT Homo sapiens cDNA clone
DEFINITION IMAGE:713888.3, similar to gb:L06797 PROBABLE G PROTEIN-COUPLED
RECEPTOR LCRI HOMOLOG (HUMAN);, mRNA sequence.

ACCESSION
NID
VERSION
KEYWORDS
SOURCE
ORGANISM

AA284857
91927416
AA284857.1 GI:1927416
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS

Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chapell,B.,
Chisoe,S., Dietrich,N., Dubuque,T., Favello,A., Gish,W.,
Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N.,
Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L.,
Roifling,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Mieg,J.,
Trevaskis,E., Underwood,K., Wohlmann,P., Waterston,R., Wilson,R.
and Marra,M.

TITLE
JOURNAL
MEDLINE
COMMENT

Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)
97044478
On May 9, 1995 this sequence version replaced gi:802981.

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Insert Length: 561 Std Error: 0.00
Seq primer: -41m13 fwd. ET from Amersham
High quality sequence stop: 294.

FEATURES

source

1. 454
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:713888"
/clone_lib="Soares ovary tumor NbhOT"
/sex="Female"
/tissue_type="ovarian tumor"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: ovary; Vector: pT73D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - Oligo(dT) primer [5',
TGTACCAATGCTGAGGAGCGCGCGCTTTTCTTTTCTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library constructed by Bento Soares and
M.Fatima Bonaldo."

BASE COUNT 159 a 95 c 66 g 134 t
ORIGIN

Query Match 19.7%; Score 342; DB 31; Length 454;
Best Local Similarity 100.0%; Pred. No. 3.9e-138;
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1346 GGCAGACCTGTGGCCAAAGTTCTTATGTTGCTGTATGTCGCGGTGAGACTGTAGAAG 1405
|||||
Db 348 GGCAGACCTGTGGCCAAAGTTCTTATGTTGCTGTATGTCGCGGTGAGACTGTAGAAG 289

QY 1406 GGAACGTGAACATTCAGAGCGGTGATGATCAGCTAAGCTAGAAATGATCCAGCTG 1465
 |||||||
 Db 288 GGAACGTGAACATTCAGAGCGGTGATGATCAGCTAAGCTAGAAATGATCCAGCTG 229
 |||||||
 QY 1466 TTATGATGAATATCTCTCCATTCCTGCGGTGGAAGCTTTTCTGTTCTTAAGAGCTGAT 1525
 |||||||
 Db 228 TTATGATGAATATCTCTCCATTCCTGCGGTGGAAGCTTTTCTGTTCTTAAGAGCTGAT 169
 |||||||
 QY 1526 TTGCTGTAGAGATGGCAGCTATATACCAAGCCCAAGTGTATAGAAATGCTGTTT 1585
 |||||||
 Db 168 TTGCTGTAGAGATGGCAGCTATATACCAAGCCCAAGTGTATAGAAATGCTGTTT 109
 |||||||
 QY 1586 TCAGTTTCAGAGATGGGTGATTTTCAGCAGCTACAGTGTACAGTCTGTTATTAAGTTGT 1645
 |||||||
 Db 108 TCAGTTTCAGAGATGGGTGATTTTCAGCAGCTACAGTGTACAGTCTGTTATTAAGTTGT 49
 |||||||
 QY 1646 TAATAAAGTACATGTTAACCTAATAAAAAAAAAAAAAAAAAAAAA 1687
 |||||||
 Db 48 TAATAAAGTACATGTTAACCTAATAAAAAAAAAAAAAAAAAAAAA 7
 |||||||
 RESULT 37
 AA284569 494 bp mRNA EST 08-AUG-1997
 LOCUS 2122e05.1 Soares ovary tumor NbHOT Homo sapiens cDNA clone
 DEFINITION IMAGE:713888.5' similar to gb:L06797 PROBABLE G PROTEIN-COUPLED
 RECEPTOR LCR1 HOMOLOG (HUMAN);, mRNA sequence.
 AA284569 g1927533
 NID A284569.1 GI:1927533
 VERSION EST.
 KEYWORDS human.
 SOURCE Homo sapiens
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 494)
 Hallier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
 Chissoe,S., Dietrich,N., Dubuque,T., Favello,A., Gish,W.,
 Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N.,
 Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L.,
 Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Mieg,J.,
 Trevasakis,E., Underwood,K., Wohlmann,P., Waterston,R., Wilson,R.
 and Maria,M.
 Generation and analysis of 280,000 human expressed sequence tags
 Genome Res. 6 (9), 807-828 (1996)
 JOURNAL MEDLINE 97044478
 COMMENT On May 9, 1995 this sequence version replaced gi:804201.
 TITLE Contact: Wilson RK
 JOURNAL Washington University School of Medicine
 MEDLINE 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 COMMENT Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 Insert length: 561 Std Error: 0.00
 Seq primer: -28ml3 revz ET from Amersham
 High quality sequence stop: 383.
 Location/Qualifiers
 1..494
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="16"
 /clone="IMAGE:713888"
 /clone_lib="Soares ovary tumor NbHOT"
 /sex="Female"
 /tissue_type="ovarian tumor"
 /lab_host="DH10B (ampicillin resistant)"
 /note="Organ: ovary; Vector: p713d (Pharmacia) with a
 modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
 strand cDNA was primed with a Not I - oligo(dT) primer [5'
 TGTTACCAATCTGAGTGGAGCGCGCGGTTTCTTTTCTTTTCTTTT 3'],
 double-stranded cDNA was size selected, ligated to Eco RI
 adapters (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of a modified p713 vector
 (Pharmacia). Library constructed by Bento Soares and
 M.Felina Bonaldo."

BASE COUNT 134 a 68 c 105 g 187 t
 ORIGIN
 Query Match 18.9%; Score 328; DB 31; Length 494;
 Best Local Similarity 99.7%; Pred. No. 4,4e-132;
 Matches 378; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1189 TTTTCTTAACTAGTACATTTTTCAGATATAAAGACTGACCAATATGTACAGTTTAT 1248
 |||||||
 Db 1 TTTTCTTAACTAGTACATTTTTCAGATATAAAGACTGACCAATATGTACAGTTTAT 60
 |||||||
 QY 1249 TCGTTGTGGATTTTGTCTGTGTTCTTTAGTTTGTGAAGTTTAACTATT 1308
 |||||||
 Db 61 TCGTTGTGGATTTTGTCTGTGTTCTTTAGTTTGTGAAGTTTAACTATT 120
 |||||||
 QY 1309 ATATAATTTTCTTGTTCATATGATGTGTCTAGGAGGACCTGTGCCAAGTTCT 1368
 |||||||
 Db 121 ATATAATTTTCTTGTTCATATGATGTGTCTAGGAGGACCTGTGCCAAGTTCT 180
 |||||||
 QY 1369 TAGTGTCTGATGTCTCGGTGAGGACTGTGAAAAGGAACTGAACATTCAGAGCGTG 1428
 |||||||
 Db 181 TAGTGTCTGATGTCTCGGTGAGGACTGTGAAAAGGAACTGAACATTCAGAGCGTG 240
 |||||||
 QY 1429 TAGTAATCAGCTAAGCTAGAAATGATCCAGCTGTTATGATATATCTCTCA 1488
 |||||||
 Db 241 TAGTAATCAGCTAAGCTAGAAATGATCCAGCTGTTATGATATATCTCTCA 300
 |||||||
 QY 1489 TTCCCGTGAAGCTTTTTCCTGTTCTTAAGAGCTATTTTGTCTAGAGATGCGACTTA 1548
 |||||||
 Db 301 TTCCCGTGAAGCTTTTTCCTGTTCTTAAGAGCTATTTTGTCTAGAGATGCGACTTA 360
 |||||||
 QY 1549 TAACCAAGCCCAAGTGG 1567
 |||||||
 Db 361 TAACCAAGCCCAAGTGG 379
 |||||||
 RESULT 38
 AA501992 308 bp mRNA EST 19-AUG-1997
 LOCUS nes1h08.s1 NCI-CGAP Co3 Homo sapiens cDNA clone IMAGE:900927.3
 DEFINITION similar to gb:L06797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCR1
 HOMOLOG (HUMAN);, mRNA sequence.
 AA501992
 NID 92236959
 VERSION AA501992.1 GI:2236959
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 308)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 JOURNAL On Sep 12, 1996 this sequence version replaced gi:1398118.
 COMMENT
 TITLE Contact: Robert Strausberg, Ph.D.
 JOURNAL Tel: (301) 496-1550
 MEDLINE Email: Robert_Strausberg@nih.gov
 COMMENT Tissue Procurement: Elias Campo, M.D., Michael R. Emmert-Buck,
 M.D., Ph.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D.
 CDNA Library Arraying: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:

www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 418 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham.

Location/Qualifiers
1. .308

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:900927"
/clone_1b="NCI_CGAP_CO3"
/sex="pooled"
/tissue_type="colon"
/lab_host="DH10B"

```

/note="vector:pr713-pac (Pharmacia) with a modified polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from 12 pooled bulk tumor samples and primed with a Not I- oligodT primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pr713 vector. Library went through one round of

BASE COUNT	100 a	69 c	52 g	87 t
ORIGIN				

Query Match	17.7%;	Score 308;	DB 34;	Length 308;
Best Local Similarity	100.0%;	Pred. No. 2,3e-123;		
Matches 308;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy 1361 CAAGTTCCTTAGTTGCTGTANGTCTCGTGAGACGTGAAAAAGGGAACTGAACATTCC 1420
|||||
|||
Db 308 CAGTCTCTTAGTGTGCTGTANGTCTCGTGAGACGTGAAAAAGGGAACCTGAACATTCC 249

QY 1421 AGAGCGTGTAGTGAATCACGTAAGCTAGAAATGATCCCACTGTTATGCATAGATTA 1489
|||||
|||
Db 248 AGAGCGTGTAGTGAATCACGTAAGCTAGAAATGATCCCACTGTTATGCATAGATTA 189

Dy 1481 TCCTCCATTCGCCGTGGAACGTTTTTCTTGTTAAAGCGTAITTTGCCTGTAGAACAAT 1540
|||||
|||
Dδ 1488 TCTTCCATTCCGGGGAACGTTTTTCTGTTCTTAAGCAGTAITTTGCTGTGAAGAAT 129

Qy 1541 GGCACCTATACCAAGCCCAAGTGTATAGAAATGCTGTTTTTCAGTTTCAGGAGT 1600
|||||
Dh 128 GGACCTTATACCAAGCCCAAGTGTATAGAAATGCTGTTTTTCAGTTTCAGGAGT 69

Dh 68 GGGTGGATTTCAGCACCCTACAGTGTAAGTCTGTATTAAGTTGTTAATAAAAGTACATG 9
Qy 1601 GGGTGGATTTCAGCACCTACAGTGTAAGTCTGTATTAAGTTGTTAATAAAAGTACATG 1666

QY	1661	TTAACTT	1668
Db	8	TTAACTT	1

RESULT	39					
A1610565/c						
LOCUS						
DEFINITION	A1610565	361 bp	MRNA	EST	21-APR-1999	
	tp41d08.x1	NCI-CGAP_U14	Hom	sapiens	CDNA clone	IMAGE:2190351 3
	similar to gb:106797	PROBABE G	PROTEIN-COUPLED	RECEPTOR	LCRI	
	HOMOLOG (HUMAN):					mrna sequence.

ACCESSION	A1610565	
NID	94619732	
VERSION	A1610565.1	GI:4619732
KEYWORDS	EST.	
SOURCE	human.	

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 361)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP)
Tumor Gene Index
Journal Unpublished (1997)

COMMENT

On Mar 10, 1998 this sequence version replaced gi:2949515.

Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emlali: koderc@caluabio.org

Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by

Clone distribution: NCI-CGAP clone distribution

found through the I.M.A.G.E. Consortium/LLN

www-bio.lnl.gov/bbrp/image/image.html

Seq primer: -40bp from Gibco

High quality sequence stop: 219.

Location/Qualifiers

FEATURES
SOURCE

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2190351"
/clone_id="MCI_CGAP_Ur4"
/tissue-type="serous papillary carcinoma, high grade, 2
pooled tumors"

```

```

/2462_086...
/note="Organ: uterus; Vector: pCMV-Sport6; Site_1: SalI;
Site_2: NotI; cloned unidirectionally. primer: Oligo dt
Average insert size 1.48 kb. Life Technologies catalog #
11542-016"

```

BASE COUNT	120 a	82 c	58 g	101 c
ORIGIN				

Query Match	17.2%	Score 299	DB 48	Length 361
Best Local Similarity	100.0%	Pred. No. 1.7e+19		
Matches 299; Conservative	0	Mismatches 0	Indels 0	Gaps 0

Qy 1373 TGCCTATGCTCCGAGTAGACCTGAGAAAAGGAGACGACCATTCAGAGGAGTAGG
 |||||
 Db 299 TGCCTATGCTCCGAGTAGACCTGAGAAAAGGAGACTGACCATTCAGAGGCTGAGT 240

QY 1433 GAATCACGTAAGCTGAAATGATCCCGACGCTTATGCAATGATATCTCCATTC 1492
|||||
|||
Db 239 GAATCAGCTAAAGCTAGAAATGATCCCGACGCTTATGCAATGATATCTCCATTC 180

```
OY      1493 CGTGAACGTTTTTCCTGTTTTAAGACGTATTTTGCTGAGAAGATGGCACTATAAC   1552
        |||||
DB      179  CGTGAACGTTTTTCCTGTTTTAAGACGTATTTCCTGAGAAGATGGCACCITATAAC   120
```

155 CAAAGCCAAAGTGTATAGAAATGCTGTTTTCACGTTTCAGGAGTGGTTCATTTC 60
119 CAAAGCCAAAGTGTATAGAAATGCTGTTTTCACGTTTCAGGAGTGGTTCATTTC 60

QY 1013 GACCCAGGATGGCTGCATCGCAGTGTTCAGTTAATAAACTTAATT
Db 59 GCACCTACAGTGACAGTCTGTATTAAGTTGTTAATAAAAGTACTGA

LOCUS	DEFINITION	RECEPTOR	LCRI	HOMOLOG	(HUMAN):	EST	16-FEB-1995
T62636	407 bp	mRNA					
LOCUS	T62636	407 bp	mRNA	EST	16-FEB-1995		
DEFINITION	yc03fil.s1	Stratagene lung (19937210)	Homo sapiens	CDNA clone			
IMAGE:79659	3'	similar to gb1.06797	PROBABLE G	PROTEIN-COUPLE			
RECEPTOR	LCRI	HOMOLOG	(HUMAN):	mRNA sequence.			

ACCESSION	T62636	
NID	9666293	
VERSION	T62636.1	GI:666293
KEYWORDS	EST.	
SOURCE	human.	

ORGANISM	REFERENCE
Homo sapiens	1 (bases 1 to 407)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;	
Eukaryota; Primates; Catarrhini; Hominae; Homo.	

AUTHORS Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiappelli, B., Chissole, S., Dietrich, N., Dubuque, T., Favellio, A., Gish, M., Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Madfis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, J., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevisan, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R., and Marra, M.

TITLE Generation and analysis of 280,000 human expressed sequence tags

JOURNAL Genome Res. 6 (9), 807-828 (1996)

MEDLINE 97044478

COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 333
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: -21m13
High quality sequence stop: 333.
Location/Qualifiers
1. 407
/organism="Homo sapiens"
/db_xref="GDB:483246"
/db_xref="taxon:9606"
/clone="IMAGE:79629"
/clone_lib="Stratagene lung (#937210)"
/sex="male"
/dev_stage="72 years"
/lab_host="SOLR cells (kanamycin resistant)"
/note="Organ: lung; Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dt. normal lung. Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGCGACGAG 3' -3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTT 3'."

BASE COUNT 141 a 88 c 65 g 112 t 1 others

ORIGIN

Query Match 17.0%; Score 298; DB 21; Length 407;
Best Local Similarity 100.0%; Pred. No. 4.5e-119;
Matches 298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1367 CTTACTTGTCTGATGTCCTGCTGAGACTGTAGAAAAGGGAAGTGAACATTCACAGCG 1426
|||||
Db 306 CTTAGTTCTGTATGTCCTGCTGAGACTGTAGAAAAGGGAAGTGAACATTCACAGCG 247

QY 1427 TGTAGTGAATCAAGTGAAGCTGAAGTATGATCCCAAGCTTTATGATGATATCTCTC 1486
|||||
Db 246 TGTATGATCAAGTGAAGCTGAAGTATGATCCCAAGCTTTATGATGATATCTCTC 187

QY 1487 CATTCGCTGAGAGCTTTTCTCTTCTTAAAGCGTGAATTTGCTGTAGAGATGGCACT 1546
|||||
Db 186 CATTCGCTGAGAGCTTTTCTCTTCTTAAAGCGTGAATTTGCTGTAGAGATGGCACT 127

QY 1547 TATAACCAAGCCCAAGTGTATGAAGTGTGTTTTCGTTTTCAGAGTGGGTTG 1606
|||||
Db 126 TATAACCAAGCCCAAGTGTATGAAGTGTGTTTTCGTTTTCAGAGTGGGTTG 67

QY 1607 ATTTCAGACCTACAGTGTATGATGTTTATTAAGTATTAAGTACATGTTAA 1664
|||||
Db 66 ATTTCAGACCTACAGTGTATGATGTTTATTAAGTATTAAGTACATGTTAA 9

RESULT 41
AA621854 342 bp mRNA EST 21-OCT-1997
LOCUS nq19901.s1 NCI CGAP, Thy1 Homo sapiens cDNA clone IMAGE:1144368
DEFINITION similar to gb:106797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCR1
HOMOLOG (HUMAN):, mRNA sequence.

ACCESSION AA621854
NID 92525730
VERSION AA621854.1 GI:2525730
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 342)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
TITLE Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1407054.

CONTACT: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: David B. Kriman, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: www.bio.lnl.gov/db/ftp/image/image.html

FEATURES
SOURCE
1. 342
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1144368"
/clone_lib="NCI-CGAP, Thy1"
/tissue_type="thyroid"
/lab_host="DH10B"
/note="Vector: pAMP10; mRNA made from invasive thyroid tumor, cDNA made by oligo-dt priming. Non-directionally cloned. Size-selected on agarose gel, average insert size 600 bp. Reference: Krizman et al. (1996) Cancer Research 56:5380-5383."

BASE COUNT 78 a 105 c 73 g 86 t

ORIGIN

Query Match 17.0%; Score 296; DB 36; Length 342;
Best Local Similarity 100.0%; Pred. No. 3.4e-118;
Matches 296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 778 AAGGGCCACCCGAAGGCGCCCTCAAGACCAAGTATCTATCTGCTTTCTTC 837
|||||
Db 12 AAGGGCCACCCGAAGGCGCCCTCAAGACCAAGTATCTATCTGCTTTCTTC 71

QY 838 GCGCTTGGCTGCTTACTACATTTGGATGACATCGATCCTTCATCCTGGAATC 897
|||||
Db 72 GCGCTTGGCTGCTTACTACATTTGGATGACATCGATCCTTCATCCTGGAATC 131

QY 898 ATGANGCAAGGGTGTGAGTTGAGAACACTGTGCACANGTGAATTCATCACGAGGCC 957
|||||
Db 132 ATGANGCAAGGGTGTGAGTTGAGAACACTGTGCACANGTGAATTCATCACGAGGCC 191

QY 958 CTAGCTTCTTCACATGTGTGTCGAACCCATCCCTATGCTTCTTGGAGCCAAATTG 1017
|||||
Db 192 CTAGCTTCTTCACATGTGTGTCGAACCCATCCCTATGCTTCTTGGAGCCAAATTG 251

QY 1018 AAACCTCTGCCACGACGACCTCACTCTGTGAGCAGAGGGTCCAGCCTCAAGT 1073
|||||
Db 252 AAACCTCTGCCACGACGACCTCACTCTGTGAGCAGAGGGTCCAGCCTCAAGT 307

RESULT 42

Thu Aug 12 09:17:07 1999

us-09-104-063-3.1.txt

Page 34

AA386001 438 bp mRNA EST 21-APR-1997
LOCUS AA386001
DEFINITION E5199759 Pancreas tumor I Homo sapiens cDNA 5' end similar to
guanine nucleotide-binding protein HM89, mRNA sequence.
ACCESSION AA386001
VERSION 92038338
KEYWORDS GI:2038338
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 438)
AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fulcher,R.A.,
Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man,Wai,C.,
White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man,Wai,C.,
Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr.,
Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaro,S.M., Merrick,D.M.,
Moreno-Palancas,R.F., McDonald,L.A., Nguyen,D.T., Shibley,R.,
Phillips,C.A., Ryder,S.E., Scott,J.L., Sauder,D.M., Shibley,R.,
Small,K.V., Spriggs,T.A., Uterback,T.R., Weidman,J.F., Li,Y.,
Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,G.A.,
Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H.,
Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
Dillon,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
Fraser,C.M., and Venter,J.C.
TITLE Initial assessment of human gene diversity and expression patterns
based upon 83 million nucleotides of cDNA sequence
JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)
MEDLINE 96026280
COMMENT On Dec 30, 1996 this sequence version replaced gi:1530938.
Other ESTs: E5199758 THC166975
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@igr.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (<http://www.tigr.org/tldb/hgi/hgi.html>)
Seq primer: M13 Reverse.
FEATURES
Location/Qualifiers
1..438
/organism="Homo sapiens"
/db_xref="ATCC (inhost):190162"
/db_xref="taxon:9606"
/clone_lib="Pancreas tumor I"
/dev_stage="adult"
/note="Organ: pancreas; Vector: pBluescript SK-; Site_1:
EcoRI; Site_2: XhoI"
BASE COUNT 104 a 104 c 83 g 145 t 2 others
ORIGIN
Query Match 17.0%; Score 295; DB 33; Length 438;
Best Local Similarity 100.0%; Pred. No. 8,7e-118;
Matches 295; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 847 CTGCTACTACATGGGATGAGCATGACTCTTCACTCTCTGGAATCATCAAGCAA 906
DB 1 CTGCTACTACATGGGATGAGCATGACTCTTCACTCTCTGGAATCATCAAGCAA 60
OY 907 GGGTGTAGTTTGAACACCTGTGCAAGTTCATCATCAGAGAGCCCTAGCTTTC 966
DB 61 GGGTGTAGTTTGAACACCTGTGCAAGTTCATCATCAGAGAGCCCTAGCTTTC 120
OY 967 TTCCAGTGTGTGAGACCCATCTATGCTTCTTGGAGCCAAATTAAACCTCT 1026

|||||
DB 121 TTCCAGTGTGTGAGACCCATCTATGCTTCTTGGAGCCAAATTAAACCTCT 180
OY 1027 GCCCAGCAGCAGCTACCTCTGTGAGCAGAGGGTTCACCCCTCAAGATCTCTCCAAAGGA 1086
DB 181 GCCCAGCAGCAGCTACCTCTGTGAGCAGAGGGTTCACCCCTCAAGATCTCTCCAAAGGA 240
OY 1087 AAGCAGGTGACATTCATCTGTTCCACTGAGTGTGAGTCTTCAAGTTTTCAC 1141
DB 241 AAGCAGGTGACATTCATCTGTTCCACTGAGTGTGAGTCTTCAAGTTTTCAC 295
RESULT 43
AA354393 456 bp mRNA EST 21-APR-1997
LOCUS AA354393
DEFINITION E512864 Jurkat T-cells V Homo sapiens cDNA 5' end similar to
guanine nucleotide-binding protein HM89, mRNA sequence.
ACCESSION AA354393
VERSION 92006897
KEYWORDS GI:2006897
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 456)
AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fulcher,R.A.,
Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man,Wai,C.,
Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr.,
Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaro,S.M., Merrick,D.M.,
Moreno-Palancas,R.F., McDonald,L.A., Nguyen,D.T., Shibley,R.,
Phillips,C.A., Ryder,S.E., Scott,J.L., Sauder,D.M., Shibley,R.,
Small,K.V., Spriggs,T.A., Uterback,T.R., Weidman,J.F., Li,Y.,
Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,G.A.,
Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H.,
Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
Dillon,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
Fraser,C.M., and Venter,J.C.
TITLE Initial assessment of human gene diversity and expression patterns
based upon 83 million nucleotides of cDNA sequence
JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)
MEDLINE 96026280
COMMENT On May 9, 1995 this sequence version replaced gi:802440.
Other ESTs: THC166975
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@igr.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (<http://www.tigr.org/tldb/hgi/hgi.html>)
Seq primer: M13 Reverse.
FEATURES
Location/Qualifiers
1..456
/organism="Homo sapiens"
/db_xref="ATCC (inhost):154827"
/db_xref="taxon:9606"
/map="12p"
/clone_lib="Jurkat T-cells V"
/cell_type="T-lymphocyte"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT 116 a 121 c 105 g 111 t 3 others
ORIGIN

```
Query Match          16.9%; Score 293; DB 32; Length 456;
Best Local Similarity 100.0%; Pred. No. 6,3e-117;
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 59 TACCACCCGATGGAGAACGCGGTTACCATGAGGGGATCATATATACACTTCAG 118
    |||
DB 33 TACCACCCGATGGAGAACGCGGTTACCATGAGGGGATCATATATACACTTCAG 92
OY 119 ATATACTACACCGAGAAATGGCTCAGGGAGCTATGACTCATGAGAAACCCGTGTC 178
    |||
DB 93 ATATACTACACCGAGAAATGGCTCAGGGAGCTATGACTCATGAGAAACCCGTGTC 152
OY 179 GTGAAGAAATGCTAATTTCAATAAATCTCTGCGCCACCATCTACTCATCTCT 238
    |||
DB 153 GTGAAGAAATGCTAATTTCAATAAATCTCTGCGCCACCATCTACTCATCTCT 212
OY 239 TAACGGCATTTGGGCAATGATGTTGTCATCTCTGTCATGGTTTACCAAGAACTGA 298
    |||
DB 213 TAACGGCATTTGGGCAATGATGTTGTCATCTCTGTCATGGTTTACCAAGAACTGA 272
OY 299 GAAGCATGACGCAAGTACAGGCTGACCTGTGAGTGGCCGACCTCTCTT 351
    |||
DB 273 GAAGCATGACGCAAGTACAGGCTGACCTGTGAGTGGCCGACCTCTCTT 325

RESULT 44
LOCUS AA747545 555 bp mRNA EST 17-MAR-1999
DEFINITION nx85d08.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1269039 3'
            similar to gb:U06797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCRI
            HOMOLOG (HUMAN);, mRNA sequence.
ACCESSION AA747545
VERSION 22787503
KEYWORDS AA747545.1 GI:2787503
SOURCE EST.
ORGANISM human.
            Homo sapiens
            Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Mammalia:
            Eutheria: Primates: Catarrhini: Homiidae: Homo.
REFERENCE 1 (bases 1 to 555)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
JOURNAL Tumor Gene Index
COMMENT Unpublished (1997)
On May 18, 1995 this sequence version replaced gi:811242.

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
Ph.D., Gerald Marti, M.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bhrp/image/image.html

Insert length: 747 Std Error: 0.00
Seq primer: -40m3 fwd. ET from Amerisham
High quality sequence stop: 332.
Location/Qualifiers
1. 555
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="20"
/clone="IMAGE:1269039"
/clone_lib="NCI_CGAP_GCB1"
/tissue_type="germinal center B cell"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
```

```
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20-, IgD-),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
primed with a Not I - oligo(dT) primer
[5'-GGTACCAATCGAGAGTGGAGCGGCCGCCCATTTTCTTTTCTTTT-
3']. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT      210 a      108 c      84 g      153 t
ORIGIN

Query Match          16.8%; Score 292; DB 38; Length 555;
Best Local Similarity 99.8%; Pred. No. 1.6e-116;
Matches 412; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY 1266 TCTTGTTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1325
    |||
DB 413 TCTTGTTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 354
OY 1326 TTCAATATGATGTGTGTGTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1385
    |||
DB 353 TTCAATATGATGTGTGTGTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 294
OY 1386 GTGTGAGGAGCTGTAGAAAAGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1445
    |||
DB 293 GTGTGAGGAGCTGTAGAAAAGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 234
OY 1446 CTAGAAATATATCCCGAGCTGTATGATGATGATGATGATGATGATGATGATGAT 1504
    |||
DB 233 CTAGAAATATATCCCGAGCTGTATGATGATGATGATGATGATGATGATGATGAT 174
OY 1505 TTCTGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1564
    |||
DB 173 TTCTGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 114
OY 1565 TGGTATAGAAATGCTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1624
    |||
DB 113 TGGTATAGAAATGCTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 54
OY 1625 TACAGTCTGTATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1677
    |||
DB 53 TACAGTCTGTATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1

RESULT 45
LOCUS AI583169 524 bp mRNA EST 06-APR-1999
DEFINITION tr98d12.x1 NCI_CGAP_Pan1 Homo sapiens cDNA clone IMAGE:2227127 3'
            similar to gb:U06797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCRI
            HOMOLOG (HUMAN);, mRNA sequence.
ACCESSION AI583169
VERSION 94569066
KEYWORDS AI583169.1 GI:4569066
SOURCE EST.
ORGANISM human.
            Homo sapiens
            Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Mammalia:
            Eutheria: Primates: Catarrhini: Homiidae: Homo.
REFERENCE 1 (bases 1 to 524)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
JOURNAL Tumor Gene Index
COMMENT Unpublished (1997)
On May 18, 1998 this sequence version replaced gi:3136741.

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
```


RESULT 47
AI582682/c 277 bp mRNA EST 06-APR-1999
LOCUS tnl5d04.x1 NCI-CGAP_Brn25 Homo sapiens cDNA clone IMAGE:2167783 3'
DEFINITION similar to gp:L06797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCRI
HOMOLOG (HUMAN); mRNA sequence.
ACCESSION AI582682
NID g4568579
VERSION AI582682.1 GI:4568579
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 277)
NCI/NINDS-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute / National Institute of Neurological
Disorders and Stroke, Brain Tumor Genome Anatomy Project
(CGAP/BI GAP), Tumor Gene Index
Unpublished (1998)
JOURNAL On May 7, 1998 this sequence version replaced gi:3121491.
COMMENT
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www.bio.lnl.gov/dbtrp/image/image.html
Seq primer: -40UP from Gibco.
Location/Qualifiers
1..277
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2167783"
/clone_lib="NCI-CGAP_Brn25"
/tissue_type="anaplastic oligodendroglioma"
/lab_host="DH10B"
/note="Organ: brain; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; Site1: Not I; Site2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTATCCATCTGAAGTGGAGCGCGCCGATGATGTTTCTTTTCTTTTCTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT73 vector.
Library is normalized, and was constructed by Bento
Soares and M.Fatima Bonaldo."

BASE COUNT 85 a 56 c 46 g 88 t
ORIGIN

Query Match 15.9%; Score 277; DB 48; Length 277;
Best Local Similarity 100.0%; Pred. No. 6.1e-110;
Matches 277; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1580 GGTTCACGTTTCAGAGGTGGTTCAGCACCACGATGACGCTGTATTA 1639
|||||
DB 97 GGTTCACGTTTCAGAGGTGGTTCAGCACCACGATGACGCTGTATTA 38
QY 1640 AGTTGTAATAAAGTACATGTTAACTTAATAAAAAA 1676
|||||
DB 37 AGTTGTAATAAAGTACATGTTAACTTAATAAAAAA 1

RESULT 48
AA934643/c 427 bp mRNA EST 07-JUL-1998
LOCUS o071c08.s1 NCI-CGAP_GC4 Homo sapiens cDNA clone IMAGE:1571630 3'
DEFINITION similar to gp:L06797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCRI
HOMOLOG (HUMAN); mRNA sequence.
ACCESSION AA934643
NID g3091855
VERSION AA934643.1 GI:3091855
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 427)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
JOURNAL On Jan 14, 1998 this sequence version replaced gi:1797648.
COMMENT
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www.bio.lnl.gov/dbtrp/image/image.html
Insert Length: 989 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 315.
Location/Qualifiers
1..427
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1571630"
/clone_lib="NCI-CGAP_GC4"
/tissue_type="pooled germ cell tumors"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from 3 pooled
germ cell tumors, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT73
vector. Library is normalized. Library was constructed by
Bento Soares and M. Fatima Bonaldo."

BASE COUNT 146 a 87 c 67 g 125 t 2 others
ORIGIN

Query Match 15.8%; Score 275; DB 40; Length 427;
Best Local Similarity 100.0%; Pred. No. 4e-109;
Matches 275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1410 CTGAACATTCAGAGCGGTGATGATCAGCAAGCTAGAAGTATCCAGCTGTTA 1469
|||||
DB 276 CTGAACATTCAGAGCGGTGATGATCAGCAAGCTAGAAGTATCCAGCTGTTA 217

BASE COUNT	109 a	81 c	62 g	95 t
ORIGIN				
Query Match	15.88;	Score 274;	DB 22;	Length 347;
Best Local Similarity	100.0%;	Pred. No. 1.1e-108;		
Matches 274;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1383	CTCGTGGTAGACAGCTGTAGAAAANGGAACGTGAACATTCACAGACCGGTGATGATCAGCTA	1442	
Db	288	CTCGGGTAGGACGGGTAGAAAAAGGAACTGACATTCACAGACCGGTGATGATCAGCTA	229	
QY	1443	AAGCAGAGAAATGATCCCGACGCTTTATGCATAGATAAATCTCTCCATTCGCCGTGAACGT	1502	
Db	228	AAGCTAGAAATGATCCCGACGCTTTATGCATAGATAAATCTCTCCATTCGCCGTGAACGT	169	
QY	1503	TTTTCCTGTTCTTAAAGCGTGAATTTGCTGTGAGAGATGGCAGCTTATAACCAAAACCCAA	1562	
Db	168	TTTTCCTGTTCTTAAAGCGTGAATTTGCTGTGAGAGATGGCAGCTTATAACCAAAACCCAA	109	
QY	1563	ACTGCTATAGAAATCTGCTGTTTTCAGTTTCAGAGAGGGTGTGATTCAGCAGCTACAG	1622	
Db	108	ACTGCTATAGAAATCTGCTGTTTTCAGTTTCAGAGAGGGTGTGATTCAGCAGCTACAG	49	
QY	1623	TGTACAGCTGTATTATTAAGTTGTTATATAAAGTA	1656	
Db	48	TGTACAGCTGTATTATTAAGTTGTTATATAAAGTA	15	
RESULT 50				
LOCUS	A1581657	274 bp	mRNA	EST
DEFINITION	as03h04.x1 Barstead colon HPLB7 Homo sapiens CDNA clone			06-APR-1999
	IMAGE:2174071 3'	similar to gp:L06797	PROBABLE G PROTEIN-COUPLED	
ACCESSION	A1581657			
NTD	94567554			
VERSION	A1581657.1	GI:4567554		
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			

[illegible]

Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: esc@wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: -400P from Gibco.
 Location/Qualifiers
 1..274
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2174071"
 /clone_lib="Barsteed colon HPLRB7"
 /sex="male"
 /dev_stage="adult, age 25"
 /lab_host="DH10B (phage resistant)"
 /note="Organ: colon; Vector: pTR17D-Pac (Pharmacia) with a
 modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st
 strand cDNA was primed with a Not I - oligo(dT) primer [5

TGTTACGAATCTGAAGTGGAGCGGCCCTTTTTTTTTTTTTTTTTT
 3'] : double-stranded cDNA was ligated to Eco RI adaptors
 [5' AATTCACCTAGTAAT 3' and 5' ATTACTAGTG 3'], digested
 with Not I and cloned into the Not I and Eco RI sites of
 the modified pT713 vector. Library constructed by Bob
 Barstead."

BASE COUNT 85 a 58 c 46 g 85 t
 ORIGIN

Query Match 15.8%: Score 274; DB 48; Length 274;
 Best Local Similarity 100.0%; Pred. No. 1.2e-108;
 Matches 274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1400 GAAAGGGAGACTGAACATTCAGAGCGGTAGTGAATCAGTAAGCTAAGTAAGATGCC 1459
 DB 274 GAAAGGGAGACTGAACATTCAGAGCGGTAGTGAATCAGTAAGCTAAGTAAGATGCC 215
 QY 1460 CAGCTGTTATGCAATAGATATCTCTCCATTCGCCGTGAACGTTTCTCTCTTAA 1519
 DB 214 CAGCTGTTATGCAATAGATATCTCTCCATTCGCCGTGAACGTTTCTCTCTTAA 155
 QY 1520 CGTGATTTTGGTGTAGAGATGGCACTTATACCAAGGCCCAAGGTGTATAGAAATGCT 1579
 DB 154 CGTGATTTTGGTGTAGAGATGGCACTTATACCAAGGCCCAAGGTGTATAGAAATGCT 95
 QY 1580 GGTTTTCAGTTTTCAGAGTGGGTGATTTTCAGCAGCTACAGTGTACAGTCTTGATTA 1639
 DB 94 GGTTTTCAGTTTTCAGAGTGGGTGATTTTCAGCAGCTACAGTGTACAGTCTTGATTA 35
 QY 1640 AGTTGTTAATAAAGTACATGTTAACTTAAAA 1673
 DB 34 AGTTGTTAATAAAGTACATGTTAACTTAAAA 1

Search completed: August 8, 1999, 02:51:16
 Job time: 2026 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 7, 1999, 21:24:15 ; Search time 13.25 Seconds
(without alignments)
629.246 Million cell updates/sec

Title: US-09-104-063-4
Perfect score: 352
Sequence: 1 MEGISITSDNYTEMGSGD.....KRGHSSVSTESSSFHSS 352

Scoring table: OLIGO

Searched: 188963 seqs, 23686106 residues

Database: A_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	352	100.0	352	1 R68812	Human monocyte PF4
2	352	100.0	352	1 R80757	Chemokine superfamily
3	347	98.6	356	1 W97362	G-protein coupled
4	337	95.7	337	1 W48734	Human RM3 seven tr
5	273	77.6	352	1 R27792	New platelet factor
6	236	67.0	337	1 R53753	Seven transmembran
7	83	23.6	92	1 W97363	A partial CXCR4B p
8	60	17.0	359	1 W64778	A murine CXCR4B p
9	25	7.1	25	1 W34001	Fusin receptor ant
10	22	6.2	22	1 W34003	Fusin receptor ant
11	22	6.2	22	1 W34005	Fusin receptor ant
12	21	6.0	21	1 W34004	Fusin receptor ant
13	21	6.0	21	1 W34006	Fusin receptor ant
14	21	6.0	21	1 W34007	Fusin receptor ant
15	12	3.4	20	1 W34002	Fusin receptor ant
16	10	2.8	350	1 R27791	Interleukin-8 rece
17	10	2.8	355	1 R28272	Interleukin-8 rece
18	10	2.8	360	1 R28273	Sequence in a low
19	10	2.8	335	1 R33420	Human IL-8 recepto
20	10	2.8	18	1 R37903	Interleukin-8 rece
21	10	2.8	350	1 R68811	Interleukin-8 rece
22	10	2.8	367	1 R67671	Mouse oploid recep
23	10	2.8	367	1 R71968	Rat oploid recep
24	10	2.8	367	1 R74298	Mouse kappa-3 opio
25	10	2.8	367	1 R76638	Rat opioiph recept
26	10	2.8	1060	1 R70123	IL8-R type 1-GBP 1
27	10	2.8	1064	1 R70124	IL8-R type 2-GBP 1
28	10	2.8	350	1 R80756	Interleukin 8 rece
29	10	2.8	360	1 R80758	Interleukin 8 rece
30	10	2.8	360	1 R80953	Recombinant high a
31	10	2.8	355	1 R80950	Recombinant high a
32	10	2.8	350	1 R80951	Recombinant high a
33	10	2.8	358	1 R80952	Recombinant high a
34	10	2.8	367	1 W25217	Rat orphanin FQ re
35	10	2.8	367	1 W26582	Rat methadone-spec
36	10	2.8	73	1 W69988	Rodent chemokine r
37	10	2.8	10	1 W49091	Interleukin 8 type
38	10	2.8	367	1 W80549	Rat methadone-spec

ALIGNMENTS

```
RESULT 1
ID R68812 standard; Protein: 352 AA.
AC R68812:
DE 18-JUL-1995 (first entry)
DE Human monocyte PF4AR.
KW Interleukin-8 receptor; IL-8 receptor; PF4AR;
KW platelet factor superfamily receptor; monocyte; chemotactic;
KW inflammation; inflammatory disease; arthritis; emphysema; cystic;
KW fibrosis; colitis; bronchitis; meningitis; therapeutic.
OS Homo sapiens.
PN WO9428931-A.
PD 22-DEC-1994.
PF 07-JUN-1994: U06380.
PR 11-JUN-1993: US-076093.
PA (GETH ) GENENTECH INC.
PI Chuntharapai A, Hebert C, Kim KJ, Lee J;
DR WPI: 95-036114/05.
DR N-PSDB: 080521.
PT Treatment of inflammatory disorders - by administering an
PT antibody capable of binding a platelet factor 4 superfamily
PI receptor polypeptide
PS Disclosure: page 54-56; 83pp; English.
CC 2 PF4AR members were identified by probing lambda libraries from
CC human monocyte-like cell line HL-60 and human peripheral blood
CC lymphocytes using a large fragment of IL-8 receptor DNA (full
CC sequence given in Q80520). The nucleotide sequences of the 2
CC PF4ARs are given in Q80521 and Q80522, and their respective
CC amino acid sequences in R68812 and R68813.
SQ Sequence 352 AA:
```

```
Query Match 100.0%; Score 352; DB 1; Length 352;
Best Local Similarity 100.0%; Pred. NO. 0;
Matches 352; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MEGISITSDNYTEMGSGDYSKMEPCFRENANFNKIFPTIYSIFLTGIVNGIYI 60
DB 1 MEGISITSDNYTEMGSGDYSKMEPCFRENANFNKIFPTIYSIFLTGIVNGIYI 60
QY 61 LVMGQKRLRSTDKYRHLSTADLFLVITLPFAVDVANNYFENFLCKAVHYITVNL 120
DB 61 LVMGQKRLRSTDKYRHLSTADLFLVITLPFAVDVANNYFENFLCKAVHYITVNL 120
QY 121 YSSVITLAFISIDRYLATVHATNSORPKLLAEKVYVYGVWIPALLTTPDIFANVSPA 180
DB 121 YSSVITLAFISIDRYLATVHATNSORPKLLAEKVYVYGVWIPALLTTPDIFANVSPA 180
QY 181 DDRYICDRFPNDLWVYVFOFHINWGLIPGIVLLSCYCIISKLSHSGHQRKALKT 240
DB 181 DDRYICDRFPNDLWVYVFOFHINWGLIPGIVLLSCYCIISKLSHSGHQRKALKT 240
QY 241 TVIILAFACMLPYRISIDISFLLEIKQGCFFENTVHKWISITALEFFHCLNPI 300
DB 241 TVIILAFACMLPYRISIDISFLLEIKQGCFFENTVHKWISITALEFFHCLNPI 300
QY 301 LYAFGAKFTSAOHALTSVSRGSSLKILSKGRGHSVSTESSSFHSS 352
DB 301 LYAFGAKFTSAOHALTSVSRGSSLKILSKGRGHSVSTESSSFHSS 352
RESULT 2
ID R80757 standard; Protein: 352 AA.
AC R80757:
DE 26-MAR-1996 (first entry)
DE Chemokine superfamily receptor.
KW Interleukin; IL-8; inflammation; psoriasis; dermatitis;
KW rheumatoid arthritis; inflammatory bowel disease;
KW chronic lung inflammation; treatment; antibody;
KW affinity purification; detection.
OS Homo sapiens.
PN US5440021-A.
```

PD 08-AUG-1995.
 PF 29-MAR-1991: 677211.
 PR 29-MAR-1991: US-677211.
 PR 25-FEB-1994: US-202056.
 PA (CHUN/) CHUNTHARAPAI A.
 PA (HEBE/) HEBERT C.
 PA (KIM/) KIM R J.
 PA (LEEJ/) LEE J.
 PI Chuntharapai A, Hebert C, Kim KJ, Lee J:
 DR WPI: 95-283151/37.
 DR N-PSDB: 099007.
 PT New antibodies against interleukin 8 type B receptor - used to treat
 or prevent inflammation, also for detecting receptor expression and
 purification.
 PS Example 2: Columns 45-48: 62pp: English.
 CC Antibodies directed against the interleukin-8 receptor B can be used
 to treat or prevent inflammation e.g. psoriasis, dermatitis,
 rheumatoid arthritis and particularly inflammatory bowel disease and
 chronic lung inflammation. When immunoblised, these antibodies may
 be used to detect interleukin-8 receptor B expression in cells and
 tissues and for affinity purification of interleukin-8 receptor B
 from cells. This sequence is an additional chemokine superfamily
 receptor which was identified by probing lambda libraries of genomic
 DNA from a human monocyte-like cell line (U-60) and human peripheral
 blood lymphocytes using a large fragment of the interleukin-8 type
 A receptor DNA (See Q99006).
 CC Sequence 352 AA:
 SQ

Query Match 100.0%; Score 352; DB 1; Length 352;

Best local Similarity 100.0%; Pred. No. 0;

Matches 352; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEGISLYTSDNTEEMSGSDYSKPECFRENNANFKIFLPIYISIFLIGVNGVLT 60
 DB 1 MEGISLYTSDNTEEMSGSDYSKPECFRENNANFKIFLPIYISIFLIGVNGVLT 60

QY 61 LVMGYKRIKRSMTDKRLHSVADLLEFVITLPEWAVDAVAMVYGFNLCRAVIVITVN 120
 DB 61 LVMGYKRIKRSMTDKRLHSVADLLEFVITLPEWAVDAVAMVYGFNLCRAVIVITVN 120

QY 121 YSSVLLAFLISDRYLAIVHATNSOPRKLAEKVYVGVWIPALLITIPDIFANVSEA 180
 DB 121 YSSVLLAFLISDRYLAIVHATNSOPRKLAEKVYVGVWIPALLITIPDIFANVSEA 180

QY 181 DREYICDREYNDLWVYVFOHIMWGLIPGIYIISCYIIISKSHSGHOKRAKLT 240
 DB 181 DREYICDREYNDLWVYVFOHIMWGLIPGIYIISCYIIISKSHSGHOKRAKLT 240

QY 241 TVLLILAFACWLPYIYIGISIDSFILLEIKOGCEFFNTVHKWISITELAFHCCLNPI 300
 DB 241 TVLLILAFACWLPYIYIGISIDSFILLEIKOGCEFFNTVHKWISITELAFHCCLNPI 300

QY 301 LVAFILGAKFTSAQHALLTSVRSGLKILSKGRGHSVSTESSSSPHSS 352
 DB 301 LVAFILGAKFTSAQHALLTSVRSGLKILSKGRGHSVSTESSSSPHSS 352

RESULT 3
 W97362
 ID W97362 standard; Protein: 356 AA.
 AC W97362:
 DE 12-MAY-1999 (first entry)
 KW G-protein coupled receptor; polypeptide designated CXCR4B.
 KW G-protein coupled receptor; CXCR4B; human; splice variant;
 KW chemokine receptor; CXCR4; vaccine; gene therapy; HIV-1; HIV-1;
 KW cancer; pain; diabetes; obesity; anorexia; bulimia; asthma;
 KW parkinson's disease; acute heart failure; hypotension; hypertension;
 KW urinary retention; osteoporosis; angina pectoris; myocardial infarction;
 KW stroke; ulcer; allergy; benign prostatic hypertrophy; migraine;
 KW vomiting; psychotic disorder; neurodegeneration; delirium; dementia; mental retardation;
 KW schizophrenia; manic depression; delirium; dementia; mental retardation;
 KW dyskinesias; Huntingtons disease; Gilles de la Tourette syndrome.

OS Homo sapiens.
 PN EP-897980-A2.
 PD 24-FEB-1999.
 PF 07-AUG-1998: 306324.
 PR 24-JUL-1998: US-035601.
 PR 20-AUG-1997: US-056601.
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 PI Gupta SK, Pillarisetti K;
 DR WPI: 99-134643/12.
 DR N-PSDB: X15882.
 PT New G protein coupled receptor (CXCR4B) polypeptide and
 polynucleotide, human splice variant of a chemokine receptor -
 useful as diagnostic reagents and for prevention and treatment of
 HIV infection, cancer, stroke and dementia
 PS Claim 1: Page 18: 24pp: English.
 CC The present sequence represents a G-protein coupled receptor polypeptide
 designated CXCR4B, which is a human splice variant of a chemokine
 receptor CXCR4. CXCR4B polynucleotides and polypeptides are useful for
 diagnosing susceptibility to diseases by detecting mutations or
 polymorphisms in the CXCR4B gene or analysing for the presence or amount
 of CXCR4B polypeptide expressed in a patient. CXCR4B polypeptides and
 polynucleotides are also useful for screening for antagonists and
 agonists which can be used to treat conditions associated with CXCR4B
 polypeptide imbalance. CXCR4B polypeptides can be administered directly
 (as a vaccine) or via a vector (gene therapy) to prevent disease.
 CC Diseases diagnosed, prevented and treated include: HIV-1 and HIV-1
 infections; cancer; pain; diabetes; obesity; anorexia; bulimia; asthma;
 CC parkinson's disease; acute heart failure; hypotension; hypertension;
 CC urinary retention; osteoporosis; angina pectoris; myocardial infarction;
 CC stroke; ulcer; allergy; benign prostatic hypertrophy; migraine;
 CC vomiting; psychotic and neurological disorders including anxiety;
 CC schizophrenia, manic depression, delirium, dementia and severe mental
 CC retardation; and dyskinesias, such as Huntingtons disease or Gilles de
 CC la Tourette syndrome. CXCR4B polypeptides are useful for mapping genes
 CC to chromosomes, allowing gene inheritance to be studied through linkage
 CC analysis.
 CC Sequence 356 AA:
 SQ

Query Match 98.6%; Score 347; DB 1; Length 356;

Best local Similarity 100.0%; Pred. No. 0;

Matches 347; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 IYTSNDYTEEMSGSDYSKPECFRENNANFKIFLPIYISIFLIGVNGVLTVMGY 65
 DB 10 IYTSNDYTEEMSGSDYSKPECFRENNANFKIFLPIYISIFLIGVNGVLTVMGY 69

QY 66 OKRLRSMTDKRYRLHSVADLLEFVITLPEWAVDAVAMVYGFNLCRAVIVITVNLSYVL 125
 DB 70 OKRLRSMTDKRYRLHSVADLLEFVITLPEWAVDAVAMVYGFNLCRAVIVITVNLSYVL 129

QY 126 ILAFISIDRYLAIVHATNSOPRKLAEKVYVGVWIPALLITIPDIFANVSEAADRYI 185
 DB 130 ILAFISIDRYLAIVHATNSOPRKLAEKVYVGVWIPALLITIPDIFANVSEAADRYI 189

QY 186 CDRFYNDLWVYVFOHIMWGLIPGIYIISCYIIISKSHSGHOKRAKLTIVILI 245
 DB 190 CDRFYNDLWVYVFOHIMWGLIPGIYIISCYIIISKSHSGHOKRAKLTIVILI 249

QY 246 LAFACWLPYIYIGISIDSFILLEIKOGCEFFNTVHKWISITELAFHCCLNPIYAF 305
 DB 250 LAFACWLPYIYIGISIDSFILLEIKOGCEFFNTVHKWISITELAFHCCLNPIYAF 309

QY 306 GAKFTSAQHALLTSVRSGLKILSKGRGHSVSTESSSSPHSS 352
 DB 310 GAKFTSAQHALLTSVRSGLKILSKGRGHSVSTESSSSPHSS 356

RESULT 4
 W48734
 ID W48734 standard; Protein: 337 AA.
 AC W48734:
 DE 25-SEP-1998 (first entry)

DE Human RM3 seven transmembrane (7TM) receptor.
 KW V28: Placenta: seven transmembrane receptor: 7TM; signal transduction:
 KM Immunology: Inflammation: RM3.
 OS Homo sapiens.

FT	Key	Location/Qualifiers
FT	Domain	48..69
FT	Domain	/note= "Transmembrane domain 1"
FT	Domain	82..100
FT	Domain	/note= "Transmembrane domain 2"
FT	Domain	115..136
FT	Domain	/note= "Transmembrane domain 3"
FT	Domain	159..179
FT	Domain	/note= "Transmembrane domain 4"
FT	Domain	198..220
FT	Domain	/note= "Transmembrane domain 5"
FT	Domain	246..274
FT	Domain	/note= "Transmembrane domain 6"
FT	Domain	287..311
FT	Domain	/note= "Transmembrane domain 7"

PN U55759804-A.
 PD 02-JUN-1998.
 PF 17-NOV-1993: 153848.
 PR 17-NOV-1992: US-977452.
 PA (ICOS-) ICOS CORP.
 PI Godiska R, Gray PM, Schweickart VL;
 DR WPI: 98-332132/29.
 DR N-PSDB: V18357.
 PT DNA encoding V28 seven transmembrane receptor polypeptide - useful
 PT for producing recombinant polypeptide and anti-V28 antibodies, and
 PT in screening assays for V28 agonists and antagonists
 PS Example 11: Columns 93-94: 56pp: English.
 CC The present sequence represents the human RM3 protein which is encoded
 CC by the full length genomic RM3 clone (V18357) which was isolated from
 CC a human macrophage cDNA library. The invention claims for the full
 CC length V28 genomic DNA sequence (V18343) isolated from a human placenta
 CC genomic library. The V28 (M48722) and RM3 proteins are seven
 CC transmembrane (7TM) receptors which are probably involved in signal
 CC transduction. The invention also claims that cells transformed with V28
 CC DNA can be used to produce the recombinant polypeptide, to produce
 CC anti-V28 antibodies or in screening assays for V28 agonists or
 CC antagonists. The antibodies, agonists and antagonists could then be
 CC used to modulate V28 receptor-ligand binding, for e.g. in immunological
 CC and/or inflammatory events in vivo.
 SQ Sequence 337 AA:

Query Match 95.7%; Score 337; DB 1; Length 337;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 337: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 MGSDDYMSKPCFREENANFNKIFLPTIYSIFLTGIVNGVLIVMGYOKLRSMIDK 75
 DB 1 MGSDDYMSKPCFREENANFNKIFLPTIYSIFLTGIVNGVLIVMGYOKLRSMIDK 60
 OY 76 YRLHLSVADLLFVITLPFAVDAVANMYFGNFLCKAVHYITVNLSSVLLAFISLDYR 135
 DB 61 YRLHLSVADLLFVITLPFAVDAVANMYFGNFLCKAVHYITVNLSSVLLAFISLDYR 120
 OY 136 LAIYHATNSQRPRLKLAEKVYVGVWIPALLITPDIFFANVSEADRYICDRFPNDLM 195
 DB 121 LAIYHATNSQRPRLKLAEKVYVGVWIPALLITPDIFFANVSEADRYICDRFPNDLM 180
 OY 196 VVVFQFQIMWGLIPGIVILSCYCIISKLSHSGHOKRKALKTIVILIAFFACWLPY 255
 DB 181 VVVFQFQIMWGLIPGIVILSCYCIISKLSHSGHOKRKALKTIVILIAFFACWLPY 240
 OY 256 YIGSISIDFILLEIKKQCEFEENTYHKWISTEALAFHCCINPLIVFLGAKFTSAQH 315
 DB 241 YIGSISIDFILLEIKKQCEFEENTYHKWISTEALAFHCCINPLIVFLGAKFTSAQH 300
 OY 316 ALTSVRSGLSKILSKGKRGHSSVTESSSSPHSS 352
 DB 301 ALTSVRSGLSKILSKGKRGHSSVTESSSSPHSS 337

RESULT 5
 ID R27792 standard: Protein: 352 AA.
 AC R27792:
 DT 12-MAR-1993 (first entry)
 DE New platelet factor 4 receptor superfamily member PF4ARL.
 KW IL-8R; G-protein coupled receptor family: rhodopsin superfamily;
 KM pro-inflammatory cytokine: 8tr.20.15.
 OS Homo sapiens.
 PN W09217497-A.
 PD 15-OCT-1992.
 PF 23-MAR-1992: U02317.
 PR 29-MAR-1991: US-677211.
 PR 19-DEC-1991: US-810782.
 PA (GERTH) GENENTECH INC.
 PI Holmes WE, Lee J, Wood WI;
 DR WPI: 92-366191/44.
 DR N-PSDB: Q29506.
 PT Isolated human platelet factor 4 super-family receptor
 PT polypeptide and corresp. antibodies and DNA - useful as
 PT diagnostic and screening agents, and for treating inflammation or
 PT PF4AR-mediated disorders
 PS Claim 7: Fig 4: 78pp: English.
 CC The IL-8 receptor cDNA sequence was isolated (see Q29505) and a
 CC 874bp sub-fragment of the coding sequence was used as a probe to
 CC screen human cell line HL60 and human peripheral blood lymphocyte
 CC cDNA libraries. Two new gene sequences were found that are clearly
 CC related to the IL-8 receptor. One of these was contained in
 CC combined clone 8tr.20.15 and is predicted to encode an amino acid
 CC sequence which is 34% identical with both the high and low affinity
 CC IL-8 receptors. See also Q37107.
 SQ Sequence 352 AA:

Query Match 77.6%; Score 273; DB 1; Length 352;
 Best Local Similarity 100.0%; Pred. No. 1e-256;
 Matches 273: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MEGISITYSNTYEMSGSDYMSKPCFREENANFNKIFLPTIYSIFLTGIVNGLVI 60
 DB 1 MEGISITYSNTYEMSGSDYMSKPCFREENANFNKIFLPTIYSIFLTGIVNGLVI 60
 OY 61 LVMGYOKLRSMIDKYYRLHLSVADLLFVITLPFAVDAVANMYFGNFLCKAVHYITVNL 120
 DB 61 LVMGYOKLRSMIDKYYRLHLSVADLLFVITLPFAVDAVANMYFGNFLCKAVHYITVNL 120
 OY 121 YSSVILIAFISIDRYLAVHATNSQRPRLKLAEKVYVGVWIPALLITPDIFFANVSEA 180
 DB 121 YSSVILIAFISIDRYLAVHATNSQRPRLKLAEKVYVGVWIPALLITPDIFFANVSEA 180
 OY 181 DDRIYCDRFPNDLMVVVFQFQIMWGLIPGIVILSCYCIISKLSHSGHOKRKALKT 240
 DB 181 DDRIYCDRFPNDLMVVVFQFQIMWGLIPGIVILSCYCIISKLSHSGHOKRKALKT 240
 OY 241 TVIILAFACWLPYIIGSISIDFILLEIKKQ 273
 DB 241 TVIILAFACWLPYIIGSISIDFILLEIKKQ 273

RESULT 6
 ID R53753 standard: Protein: 337 AA.
 AC R53753:
 DT 07-FEB-1995 (first entry)
 DE Seven transmembrane receptor (RM3).
 KW Primer: seven transmembrane receptor; receptor: amplification: PCR;
 KM polymerase chain reaction.
 OS Homo sapiens.
 PI
 DR
 FT Key Location/Qualifiers
 FT domain 48..69
 FT domain /label= Transmembrane domain.

FT domain 82..100
 /label= Transmembrane domain.
 FT domain 115..136
 /label= Transmembrane domain.
 FT domain 159..179
 /label= Transmembrane domain.
 FT domain 198..220
 /label= Transmembrane domain.
 FT domain 246..274
 /label= Transmembrane domain.
 FT domain 287..311
 /label= Transmembrane domain.
 FT domain
 WO9412635-A.
 PD 09-JUN-1994.
 PF 17-NOV-1993; 011153.
 PR 17-NOV-1992; US-977452.
 PA (ICOS-) ICOS CORP.
 PI Godiska R Gray PW, Schweickart VL;
 DR WPI: 94-200264/24.
 N-PSDB: Q66179.
 PI DNA encoding seven transmembrane receptors - used to develop
 PI prods. for use as therapeutic or diagnostic agents for conditions
 PI involving the receptors.
 PS Example 11: Page 84-85; 100pp: English.
 CC Two primers (Q66148, Q66149) were used in a PCR reaction containing
 CC human macrophage cDNA library in plasmid pRC/CMV. When the PCR
 CC products were subjected to agarose gel electrophoresis a faint band
 CC of 180-200 base pairs was observed. Re-amplified material was
 CC digested with BamH1 and HindIII and cloned into the plasmid
 CC Bluescript SK-. Of sixteen clones sequenced, two contained a unique
 CC sequence termed RM3. Specific primers for the partial RM3 clone were
 CC used to identify a full length RM3 cDNA clone (Q66179) which
 CC encoded this seven transmembrane receptor.
 SQ Sequence 337 AA:

Query Match 67.0%; Score 236; DB 1; Length 337;
 Best Local Similarity 99.7%; Pred. No. 7e-221;
 Matches 336; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 16 MGSQDYDSMKPEPCFRENNANFNKIFLPTIYSIIFLGI VNGGLVILVMGYOKLRSMWTK 75
 DB 1 MGSQDYDSMKPEPCFRENNANFNKIFLPTIYSIIFLGI VNGGLVILVMGYOKLRSMWTK 60
 QY 76 YRLHLVADLLFVTTLPFWAVDANVWYFGNFKCAVHVITYVNLSSVLLAFISIDRY 135
 DB 61 YRLHLVADLLFVTTLPFWAVDANVWYFGNFKCAVHVITYVNLSSVLLAFISIDRY 120
 QY 136 LAIVHATNSORPKRLAEKVVYVGVWIPALDLITPDEFANVSEADRYICDRFYNDLW 195
 DB 121 LAIVHATNSORPKRLAEKVVYVGVWIPALDLITPDEFANVSEADRYICDRFYNDLW 180
 QY 196 VVWFQFOHIMWGLIPGIVILSCYCIISKLSHSGHOKRAKLTYYLLIAFACWLPY 255
 DB 181 VVWFQFOHIMWGLIPGIVILSCYCIISKLSHSGHOKRAKLTYYLLIAFACWLPY 240
 QY 256 YIGISIDSFLLEITIKOGCEEFNVHKWISTEALAFPHCLNPLVAFGAKFKTSNOH 315
 DB 241 YIGISIDSFLLEITIKOGCEEFNVHKWISTEALAFPHCLNPLVAFGAKFKTSNOH 300
 QY 316 ALTSVSRGSSLKILSKRGKGGHSSVSTESSESSPSSS 352
 DB 301 ALTSVSRGSSLKILSKRGKGGHSSVSTESSESSPSSS 337
 RESULT 7
 W97363
 ID W97363 standard; Protein: 92 AA.
 AC W97363: 12-MAY-1999 (first entry)
 DT A partial CXCR4B protein.
 DE G-protein coupled receptor; CXCR4B; human; splice variant;
 KM chemokine receptor; CXCR4; vaccine; gene therapy; HIV-1; HIV-1;

KW cancer; pain; diabetes; obesity; anorexia; bulimia; asthma;
 KW Parkinson's disease; acute heart failure; hypotension; hypertension;
 KW urinary retention; osteoporosis; angina pectoris; myocardial infarction;
 KW stroke; ulcer; allergy; benign prostatic hypertrophy; migraine;
 KW vomiting; psychotic disorder; neurological disorder; anxiety;
 KW schizophrenia; manic depression; delirium; dementia; mental retardation;
 KW dyskinesias; Huntingtons disease; Gilles de la Tourette syndrome.
 OS Homo sapiens.
 PN EP-897980-R2.
 PD 24-FEB-1998.
 PF 07-AUG-1998; 306324.
 PR 24-JUL-1998; US-056601.
 PR 20-AUG-1997; US-056601.
 PA (SMK) SMITHKLINE BEECHAM CORP.
 PI Gupta SK, Pillarsetti K;
 DR WPI: 99-134643/12.
 N-PSDB: X15883.
 DR New G protein coupled receptor (CXCR4B) polypeptide and
 PT polynucleotide, human splice variant of a chemokine receptor -
 PT useful as diagnostic reagents and for prevention and treatment of
 PT HIV infection; cancer, stroke and dementia
 PS Claim 13; Page 18; 24pp: English.
 CC The present sequence represents a partial G-protein coupled receptor
 CC designated CXCR4B, which is a human splice variant of a chemokine
 CC receptor CXCR4. CXCR4B polynucleotides and polypeptides are useful for
 CC diagnosing susceptibility to diseases by detecting mutations or
 CC polymorphisms in the CXCR4B gene or analyzing for the presence or amount
 CC of CXCR4B polypeptide expressed in a patient. CXCR4B polypeptides and
 CC polynucleotides are also useful for screening for antagonists and
 CC agonists which can be used to treat conditions associated with CXCR4B
 CC polypeptide imbalance. CXCR4B polypeptides can be administered directly
 CC (as a vaccine) or via a vector (gene therapy) to prevent disease.
 CC diseases diagnosed, prevented and treated include: HIV-1 and HIV-1
 CC infections; cancer; pain; diabetes; obesity; anorexia; bulimia; asthma;
 CC Parkinson's disease; acute heart failure; hypotension; hypertension;
 CC urinary retention; osteoporosis; angina pectoris; myocardial infarction;
 CC stroke; ulcers; allergies; benign prostatic hypertrophy; migraine;
 CC vomiting; psychotic and neurological disorders, including anxiety,
 CC schizophrenia, manic depression, delirium, dementia and severe mental
 CC retardation; and dyskinesias, such as Huntingtons disease or Gilles de
 CC la Tourette syndrome. CXCR4B polypeptides are useful for mapping genes
 CC to chromosomes, allowing gene inheritance to be studied through linkage
 CC analysis.
 SQ Sequence 92 AA:

Query Match 23.6%; Score 83; DB 1; Length 92;
 Best Local Similarity 100.0%; Pred. No. 4.1e-73;
 Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 6 IYTSNTTEEMSGSDYSMKPEPCFRENNANFNKIFLPTIYSIIFLGI VNGGLVILVMGY 65
 DB 10 IYTSNTTEEMSGSDYSMKPEPCFRENNANFNKIFLPTIYSIIFLGI VNGGLVILVMGY 69
 QY 66 OKRLRSMWTKYRLHLVADLLFV 88
 DB 70 OKRLRSMWTKYRLHLVADLLFV 92
 RESULT 8
 W64778
 ID W64778 standard; Protein: 359 AA.
 AC W64778:
 DT 20-NOV-1998 (first entry)
 DE A murine CXC chemokine receptor.
 KW Mouse; CXC chemokine receptor; pre-B cell line DW34;
 KW CXC chemokine pre-B cell stimulatory factor PBSF/SDF-1;
 KW HIV infection; screening; inhibitor; AIDS.
 OS Mus sp.
 PN W09835035-A1.
 PD 13-AUG-1998.
 PF 07-FEB-1997; J00299.
 PR 07-FEB-1997; WO-J00299.

PA (SHIO) SHIONOGI & CO LTD.
PI Iizasa H, Kishimoto T, Nagasawa T, Nakajima T, Tachibana K,
PI Yoshida N, Yoshie O;
DR WPI: 98-447232/38.
DR N-PSDB: V46370.
PT Mouse CXC chemokine receptor binding to PSF/SDF-1 pre-B cell
PT stimulatory factor - is useful for screening of potential HIV
PT infection and AIDS inhibitors
PS Claim 1: Pages 54-56; 75pp; Japanese.
CC The present sequence represents a murine CXC chemokine receptor which
CC binds to the mouse CXC chemokine pre-B cell stimulatory factor
CC PSF/SDF-1. The nucleic acid is isolated from mouse pre-B cell
CC line DM34. The receptor and cells expressing it can be used in the
CC study and mapping of the mechanism of HIV infection and in screening
CC of potential inhibitors of HIV infection and the development of AIDS.
SQ Sequence 359 AA;

Query Match 17.0%; Score 60; DB 1; Length 359;
Best Local Similarity 100.0%; Pred. No. 2.7e-50;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 197 VWFQFHIMWGLPGVILSCYCIISKLSHSGHQRALKTTVILLIAFFACWLPY 256
|||||
Db 204 VWFQFHIMWGLPGVILSCYCIISKLSHSGHQRALKTTVILLIAFFACWLPY 263

RESULT 9
W34001
ID W34001 standard; peptide: 25 AA.
AC W34001;
DT 28-MAY-1998 (first entry)
DE Fusin receptor antagonist.
KW Antagonist; integral membrane protein; inhibitor; transmembrane domain;
KW therapy; psychotic disorder; Huntington's disease; Parkinson's disease;
KW post-myocardial infarction; HIV infection; cell proliferation disorder;
KW peptic ulcer; glaucoma; asthma; renal disease; autoimmune disease;
KW antibacterial; fusin receptor.
OS Homo sapiens.
PN WO9735881-A2.
PD 02-OCT-1997.
PE 26-MAR-1997; CA0203.
PR 20-AUG-1996; US-024240.
PR 27-MAR-1996; US-014306.
PR 25-JUN-1996; US-670119.
PA (GEOR/) GEORGE S R.
PA (NGCY/) NG G Y K.
PA (ODOM/) O'DOMD B F.
PA (SEEM/) SEEMAN P.
PI George SR, Ng GYK, ODowd BF, Seeman P;
DR WPI: 97-489566/45.
PT Inhibitor of integral membrane protein - used to treat e.g.
PT schizophrenia, hypertension, viral infection, cancer etc.
PS Disclosure: Page 88; 127pp; English.
CC This sequence represents a fusin receptor antagonist, and is an
CC antagonist of the invention. The antagonists (A) are for inhibiting the
CC function of a prokaryotic or eukaryotic integral membrane protein (IMP)
CC having at least one transmembrane domain (TMD), comprises a peptide
CC including at least 4 consecutive amino acids (aa) from the sequence of
CC the TMD. (A) are used to treat or prevent disorders in mammals that
CC involve disturbances of IMP, and the same effect is achieved by inserting
CC a nucleic acid encoding (A) into the cells of the mammal. Specified
CC conditions that can be treated are schizophrenia, psychotic disorders,
CC Huntington's disease, Tourette's syndrome, hypertension, post-myocardial
CC infarction, tachyarrhythmia, human immunodeficiency virus infection,
CC AIDS, cell proliferation (cancer, psoriasis or hyperkeratosis), and drug
CC abuse. (A) can also be used to treat peptic ulcer, glaucoma, asthma,
CC renal disease, Parkinson's disease, congestive heart failure, migraine,
CC depression, obesity, diabetic neuropathy/retinopathy, autoimmune
CC diseases, graft rejection etc. (A) can be used as diuretics and
CC antibacterials. (A) may also be used to determine function of
CC orphan receptors and, when attached to a marker, for tissue imaging to
CC localise/quantify specific receptors. The transgenic animals are used to

CC test efficacy/toxicity of (A) and as models of diseases. (A) are very
CC specific and selective, against the IMP from which they are derived, so
CC should have few side effects.
SQ Sequence 25 AA;

Query Match 7.1%; Score 25; DB 1; Length 25;
Best Local Similarity 100.0%; Pred. No. 2.1e-17;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 FLPTIYSIFLTGIVNGVILVWG 64
|||||
Db 1 FLPTIYSIFLTGIVNGVILVWG 25

RESULT 10
W34003
ID W34003 standard; peptide: 22 AA.
AC W34003;
DT 28-MAY-1998 (first entry)
DE Fusin receptor antagonist.
KW Antagonist; integral membrane protein; inhibitor; transmembrane domain;
KW therapy; psychotic disorder; Huntington's disease; Parkinson's disease;
KW post-myocardial infarction; HIV infection; cell proliferation disorder;
KW peptic ulcer; glaucoma; asthma; renal disease; autoimmune disease;
KW antibacterial; fusin receptor.
OS Homo sapiens.
PN WO9735881-A2.
PD 02-OCT-1997.
PE 26-MAR-1997; CA0203.
PR 20-AUG-1996; US-024240.
PR 27-MAR-1996; US-014306.
PR 25-JUN-1996; US-670119.
PA (GEOR/) GEORGE S R.
PA (NGCY/) NG G Y K.
PA (ODOM/) O'DOMD B F.
PA (SEEM/) SEEMAN P.
PI George SR, Ng GYK, ODowd BF, Seeman P;
DR WPI: 97-489566/45.
PT Inhibitor of integral membrane protein - used to treat e.g.
PT schizophrenia, hypertension, viral infection, cancer etc.
PS Disclosure: Page 88; 127pp; English.
CC This sequence represents a fusin receptor antagonist, and is an
CC antagonist of the invention. The antagonists (A) are for inhibiting the
CC function of a prokaryotic or eukaryotic integral membrane protein (IMP)
CC having at least one transmembrane domain (TMD), comprises a peptide
CC including at least 4 consecutive amino acids (aa) from the sequence of
CC the TMD. (A) are used to treat or prevent disorders in mammals that
CC involve disturbances of IMP, and the same effect is achieved by inserting
CC a nucleic acid encoding (A) into the cells of the mammal. Specified
CC conditions that can be treated are schizophrenia, psychotic disorders,
CC Huntington's disease, Tourette's syndrome, hypertension, post-myocardial
CC infarction, tachyarrhythmia, human immunodeficiency virus infection,
CC AIDS, cell proliferation (cancer, psoriasis or hyperkeratosis), and drug
CC abuse. (A) can also be used to treat peptic ulcer, glaucoma, asthma,
CC renal disease, Parkinson's disease, congestive heart failure, migraine,
CC depression, obesity, diabetic neuropathy/retinopathy, autoimmune
CC diseases, graft rejection etc. (A) can be used as diuretics and
CC antibacterials. (A) may also be used to determine function of
CC orphan receptors and, when attached to a marker, for tissue imaging to
CC localise/quantify specific receptors. The transgenic animals are used to
CC test efficacy/toxicity of (A) and as models of diseases. (A) are very
CC specific and selective, against the IMP from which they are derived, so
CC should have few side effects.
SQ Sequence 22 AA;

Query Match 6.2%; Score 22; DB 1; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 111 AVHYITVNIYSSVLIARISL 132
|||||

Db 1 AVHIVTNLYSSVILAFISL 22

RESULT 11

W34005 standard: peptide: 22 AA.

W34005:

AC W34005: (first entry)

DT 28-MAY-1998

DE Fusin receptor antagonist.

KW Antagonist; integral membrane protein; inhibitor; transmembrane domain;

KW therapy; psychotic disorder; Huntington's disease; Parkinson's disease;

KW post-myocardial infarction; HIV infection; cell proliferation disorder;

KW peptic ulcer; glaucoma; asthma; renal disease; autoimmune disease;

OS antibacterial; fusin receptor.

OS Homo sapiens.

PN W09735881-A2.

PD 02-OCT-1997.

PF 26-MAR-1997; CA0203.

PR 20-AUG-1996; US-024240.

PR 27-MAR-1996; US-014306.

PR 25-JUN-1996; US-670119.

PA (GEOR/) GEORGE S R.

PA (NGCY/) NG G Y K.

PA (ODOW/) O'DOWD B F.

PA (SEEM/) SEEMAN P.

PI George SR, Ng GYK, ODOWD BF, Seeman P;

PI WFI: 97-489566/45.

PT Inhibitor of integral membrane protein - used to treat e.g.

PT schizophrenia, hypertension, viral infection, cancer etc.

PS Disclosure: Page 88; 127pp; English.

CC This sequence represents a fusin receptor antagonist, and is an

CC antagonist of the invention. The antagonists (A) are for inhibiting the

CC function of a prokaryotic or eukaryotic integral membrane protein (IMP)

CC having at least one transmembrane domain (TMD), comprises a peptide

CC including at least 4 consecutive amino acids (aa) from the sequence of

CC the TMD. (A) are used to treat or prevent disorders in mammals that

CC involve disturbances of IMP, and the same effect is achieved by inserting

CC a nucleic acid encoding (A) into the cells of the mammal. Specified

CC conditions that can be treated are schizophrenia, psychotic disorders,

CC Huntington's disease, Tourette's syndrome, hypertension, post-myocardial

CC infarction, tachyarrhythmia, human immunodeficiency virus infection,

CC AIDS, cell proliferation (cancer, psoriasis or hyperkeratosis), and drug

CC abuse. (A) can also be used to treat peptic ulcer, glaucoma, asthma,

CC renal disease, Parkinson's disease, congestive heart failure, migraine,

CC depression, obesity, diabetic neuropathy/retinopathy, autoimmune

CC diseases, graft rejection etc. (A) can be used as diuretics and

CC antibacterials. (A) may also be used to determine function of

CC orphan receptors and, when attached to a marker, for tissue imaging to

CC localise/quantify specific receptors. The transgenic animals are used to

CC test efficacy/toxicity of (A) and as models of diseases. (A) are very

CC specific and selective against the IMP from which they are derived, so

CC should have few side effects.

CC Sequence 22 AA:

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

KW peptic ulcer; glaucoma; asthma; renal disease; autoimmune disease;

KW antibacterial; fusin receptor.

OS Homo sapiens.

PN W09735881-A2.

PD 02-OCT-1997.

PF 26-MAR-1997; CA0203.

PR 20-AUG-1996; US-024240.

PR 27-MAR-1996; US-014306.

PR 25-JUN-1996; US-670119.

PA (GEOR/) GEORGE S R.

PA (NGCY/) NG G Y K.

PA (ODOW/) O'DOWD B F.

PA (SEEM/) SEEMAN P.

PI George SR, Ng GYK, ODOWD BF, Seeman P;

PI WFI: 97-489566/45.

PT Inhibitor of integral membrane protein - used to treat e.g.

PT schizophrenia, hypertension, viral infection, cancer etc.

PS Disclosure: Page 88; 127pp; English.

CC This sequence represents a fusin receptor antagonist, and is an

CC antagonist of the invention. The antagonists (A) are for inhibiting the

CC function of a prokaryotic or eukaryotic integral membrane protein (IMP)

CC having at least one transmembrane domain (TMD), comprises a peptide

CC including at least 4 consecutive amino acids (aa) from the sequence of

CC the TMD. (A) are used to treat or prevent disorders in mammals that

CC involve disturbances of IMP, and the same effect is achieved by inserting

CC a nucleic acid encoding (A) into the cells of the mammal. Specified

CC conditions that can be treated are schizophrenia, psychotic disorders,

CC Huntington's disease, Tourette's syndrome, hypertension, post-myocardial

CC infarction, tachyarrhythmia, human immunodeficiency virus infection,

CC AIDS, cell proliferation (cancer, psoriasis or hyperkeratosis), and drug

CC abuse. (A) can also be used to treat peptic ulcer, glaucoma, asthma,

CC renal disease, Parkinson's disease, congestive heart failure, migraine,

CC depression, obesity, diabetic neuropathy/retinopathy, autoimmune

CC diseases, graft rejection etc. (A) can be used as diuretics and

CC antibacterials. (A) may also be used to determine function of

CC orphan receptors and, when attached to a marker, for tissue imaging to

CC localise/quantify specific receptors. The transgenic animals are used to

CC test efficacy/toxicity of (A) and as models of diseases. (A) are very

CC specific and selective against the IMP from which they are derived, so

CC should have few side effects.

CC Sequence 21 AA:

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

RESULT 13

W34006 standard: peptide: 21 AA.

W34006: (first entry)

DT 28-MAY-1998

DE Fusin receptor antagonist.

KW Antagonist; integral membrane protein; inhibitor; transmembrane domain;

KW therapy; psychotic disorder; Huntington's disease; Parkinson's disease;

KW post-myocardial infarction; HIV infection; cell proliferation disorder;

KW peptic ulcer; glaucoma; asthma; renal disease; autoimmune disease;

KW antibacterial; fusin receptor.

OS Homo sapiens.

PN W09735881-A2.

PD 02-OCT-1997.

PF 26-MAR-1997; CA0203.

PR 20-AUG-1996; US-024240.

PR 27-MAR-1996; US-014306.

PR 25-JUN-1996; US-670119.

PA (GEOR/) GEORGE S R.

PA (NGCY/) NG G Y K.

PA (ODOW/) O'DOWD B F.

PI

PI

PI

PI

PA (SEEM/) SEEMAN P.
PI George SR, Ng GYK, O'Dowd BF, Seeman P;
DR WPI: 97-48956/45.
PT Inhibitor of integral membrane protein - used to treat e.g.
PT schizophrenia, hypertension, viral infection, cancer etc.
PS Disclosure: Page 88; 127pp; English.
CC This sequence represents a fusin receptor antagonist, and is an
CC antagonist of the invention. The antagonists (A) are for inhibiting the
CC function of a prokaryotic or eukaryotic integral membrane protein (IMP)
CC having at least one transmembrane domain (TMD), comprises a peptide
CC including at least 4 consecutive amino acids (aa) from the sequence of
CC the TMD. (A) are used to treat or prevent disorders in mammals that
CC involve disturbances of IMP, and the same effect is achieved by inserting
CC a nucleic acid encoding (A) into the cells of the mammal. Specified
CC conditions that can be treated are schizophrenia, psychotic disorders,
CC Huntington's disease, Tourette's syndrome, hypertension, post-myocardial
CC infarction, tachyarrhythmia, human immunodeficiency virus infection,
CC AIDS, cell proliferation (cancer, psoriasis or hyperkeratosis), and drug
CC abuse. (A) can also be used to treat peptic ulcer, glaucoma, asthma,
CC renal disease, Parkinson's disease, congestive heart failure, migraine,
CC depression, obesity, diabetic neuropathy/retinopathy, autoimmune
CC diseases, graft rejection etc. (A) can be used as diuretics and
CC antibacterials. (A) may also be used to determine function of
CC orphan receptors and, when attached to a marker, for tissue imaging to
CC localise/quantify specific receptors. The transgenic animals are used to
CC test efficacy/toxicity of (A) and as models of diseases. (A) are very
CC specific and selective, against the IMP from which they are derived, so
CC should have few side effects.
SQ Sequence 21 AA:

Query Match 5.0%; Score 21; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.3e-13;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 241 TVILIAFFACWLPIYIGISI 261
|||||
Db 1 TVILIAFFACWLPIYIGISI 21

RESULT 14
W34007
ID W34007 standard; peptide: 21 AA.
AC W34007:
DT 28-MAY-1998 (first entry)
DE Fusin receptor antagonist.
KW Antagonist; integral membrane protein; inhibitor; transmembrane domain;
KW therapy; psychotic disorder; Huntington's disease; Parkinson's disease;
KW post-myocardial infarction; HIV infection; cell proliferation disorder;
KW peptic ulcer; glaucoma; asthma; renal disease; autoimmune disease;
KW antibacterial; fusin receptor.
OS Homo sapiens.
PN WO9735881-A2.
PD 02-OCT-1997.
PF 26-MAR-1997; CA0203.
PR 20-AUG-1996; US-024240.
PR 27-MAR-1996; US-014306.
PR 25-JUN-1996; US-670119.
PA (GEOR/) GEORGE S R.
PA (NGCY/) NG G Y K.
PA (ODOW/) O'DOWD B F.
PA (SEEM/) SEEMAN P.
PI George SR, Ng GYK, O'Dowd BF, Seeman P;
DR WPI: 97-48956/45.
PT Inhibitor of integral membrane protein - used to treat e.g.
PT schizophrenia, hypertension, viral infection, cancer etc.
PS Disclosure: Page 88; 127pp; English.
CC This sequence represents a fusin receptor antagonist, and is an
CC antagonist of the invention. The antagonists (A) are for inhibiting the
CC function of a prokaryotic or eukaryotic integral membrane protein (IMP)
CC having at least one transmembrane domain (TMD), comprises a peptide
CC including at least 4 consecutive amino acids (aa) from the sequence of
CC the TMD. (A) are used to treat or prevent disorders in mammals that

CC involve disturbances of IMP, and the same effect is achieved by inserting
CC a nucleic acid encoding (A) into the cells of the mammal. Specified
CC conditions that can be treated are schizophrenia, psychotic disorders,
CC Huntington's disease, Tourette's syndrome, hypertension, post-myocardial
CC infarction, tachyarrhythmia, human immunodeficiency virus infection,
CC AIDS, cell proliferation (cancer, psoriasis or hyperkeratosis), and drug
CC abuse. (A) can also be used to treat peptic ulcer, glaucoma, asthma,
CC renal disease, Parkinson's disease, congestive heart failure, migraine,
CC depression, obesity, diabetic neuropathy/retinopathy, autoimmune
CC diseases, graft rejection etc. (A) can be used as diuretics and
CC antibacterials. (A) may also be used to determine function of
CC orphan receptors and, when attached to a marker, for tissue imaging to
CC localise/quantify specific receptors. The transgenic animals are used to
CC test efficacy/toxicity of (A) and as models of diseases. (A) are very
CC specific and selective, against the IMP from which they are derived, so
CC should have few side effects.
SQ Sequence 21 AA:

Query Match 5.0%; Score 21; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.3e-13;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 286 ITEALAFHCCLNPIYAFLG 306
|||||
Db 1 ITEALAFHCCLNPIYAFLG 21

RESULT 15
W34002
ID W34002 standard; peptide: 20 AA.
AC W34002:
DT 28-MAY-1998 (first entry)
DE Fusin receptor antagonist.
KW Antagonist; integral membrane protein; inhibitor; transmembrane domain;
KW therapy; psychotic disorder; Huntington's disease; Parkinson's disease;
KW post-myocardial infarction; HIV infection; cell proliferation disorder;
KW peptic ulcer; glaucoma; asthma; renal disease; autoimmune disease;
KW antibacterial; fusin receptor.
OS Homo sapiens.
PN WO9735881-A2.
PD 02-OCT-1997.
PF 26-MAR-1997; CA0203.
PR 20-AUG-1996; US-024240.
PR 27-MAR-1996; US-014306.
PR 25-JUN-1996; US-670119.
PA (GEOR/) GEORGE S R.
PA (NGCY/) NG G Y K.
PA (ODOW/) O'DOWD B F.
PA (SEEM/) SEEMAN P.
PI George SR, Ng GYK, O'Dowd BF, Seeman P;
DR WPI: 97-48956/45.
PT Inhibitor of integral membrane protein - used to treat e.g.
PT schizophrenia, hypertension, viral infection, cancer etc.
PS Disclosure: Page 88; 127pp; English.
CC This sequence represents a fusin receptor antagonist, and is an
CC antagonist of the invention. The antagonists (A) are for inhibiting the
CC function of a prokaryotic or eukaryotic integral membrane protein (IMP)
CC having at least one transmembrane domain (TMD), comprises a peptide
CC including at least 4 consecutive amino acids (aa) from the sequence of
CC the TMD. (A) are used to treat or prevent disorders in mammals that
CC involve disturbances of IMP, and the same effect is achieved by inserting
CC a nucleic acid encoding (A) into the cells of the mammal. Specified
CC conditions that can be treated are schizophrenia, psychotic disorders,
CC Huntington's disease, Tourette's syndrome, hypertension, post-myocardial
CC infarction, tachyarrhythmia, human immunodeficiency virus infection,
CC AIDS, cell proliferation (cancer, psoriasis or hyperkeratosis), and drug
CC abuse. (A) can also be used to treat peptic ulcer, glaucoma, asthma,
CC renal disease, Parkinson's disease, congestive heart failure, migraine,
CC depression, obesity, diabetic neuropathy/retinopathy, autoimmune
CC diseases, graft rejection etc. (A) can be used as diuretics and
CC antibacterials. (A) may also be used to determine function of
CC orphan receptors and, when attached to a marker, for tissue imaging to

CC localise/quantify specific receptors. The transgenic animals are used to
 CC test efficacy/toxicity of (A) and as models of diseases. (A) are very
 CC specific and selective, against the IMP from which they are derived, so
 CC should have few side effects.
 SQ Sequence 20 AA:

Query Match 3.4%; Score 12; DB 1; Length 20;
 Best Local Similarity 100.0%; Pred. No. 6.7e-05;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 88 VITLPMVAV 99
 9 VITLPMVAV 20

RESULT 16

R27791 ID R27791 standard; Protein; 350 AA.

AC R27791; 12-MAR-1993 (first entry)
 DE Interleukin-8 receptor.
 IL-8R; G-protein coupled receptor family; rhodopsin superfamily;
 KW pro-inflammatory cytokine.

OS Homo sapiens. Location/Qualifiers
 FH Key 40..67
 FT domain /label= transmembrane
 FT /note= "putative"
 FT 76..98
 FT domain /label= transmembrane
 FT /note= "putative"
 FT 112..133
 FT domain /label= transmembrane
 FT /note= "putative"
 FT 155..174
 FT domain /label= transmembrane
 FT /note= "putative"
 FT 204..226
 FT domain /label= transmembrane
 FT /note= "putative"
 FT 243..264
 FT domain /label= transmembrane
 FT /note= "putative"
 FT 292..312
 FT domain /label= transmembrane
 FT /note= "putative"

PN WC9217497-A.
 PD 15-OCT-1992.
 PF 23-MAR-1992; U02317.
 PR 23-MAR-1991; US-677211.
 PR 19-DEC-1991; US-810782.
 PA (GUTH) GENENTECH INC.
 PI Holmes WE, Lee J, Wood WI;
 DR WPI: 92-366191/44.
 DR N-PSDB: Q29505.
 PT Isolated human platelet factor 4 super-family receptor
 PT polypeptide and corresp. antibodies and DNA - useful as
 PT diagnostic and screening agents, and for treating inflammation or
 PT PF4AR-mediated disorders
 PS Claim 7; Fig 2; 78pp; English.
 CC A cDNA library constructed from human neutrophil mRNA in the mammalian
 CC expression vector pRK5B was transfected into COS-7 cells as pools of
 CC 2500 clones. One positive pool from the first 58 transfections was
 CC partitioned into smaller pools until a pure clone (pRK5B.1161.1)
 CC was obtained. The ORF encodes a protein of 350 amino acids which
 CC shares several features with the G-protein coupled receptors of the
 CC rhodopsin superfamily, including 7 hydrophobic (transmembrane)
 CC domains. The IL-8 receptor is a preferred PF4AR superfamily member
 CC of the invention. See also Q29506 and Q37107.
 SQ Sequence 350 AA:

Query Match 2.8%; Score 10; DB 1; Length 350;
 Best Local Similarity 100.0%; Pred. No. 0.075;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 133 DRYLAIVHAT 142
 134 DRYLAIVHAT 143

RESULT 17

R28272 ID R28272 standard; Protein; 355 AA.

AC R28272; 04-APR-1993 (first entry)
 DE Sequence in a high affinity recombinant rabbit interleukin-8
 DE (IL-8) receptor polypeptide in F3R.
 KW IL-8 receptor polypeptide; G-protein-coupled receptor.
 OS Oryctolagus cuniculus.

PN W09218641-A.
 PD 29-OCT-1992.
 PF 10-APR-1992; U02977.
 PR 10-APR-1991; US-685101.
 PR 09-JUL-1991; US-726606.
 PR 09-DEC-1991; US-803842.
 PA (REPK) REPLIGEN CORP.
 PA (UYBO-) UNIV BOSTON.
 PI Navarro J, Thomas KM, Wilt DP;
 DR WPI: 92-382123/46.
 DR N-PSDB: Q30011.
 PT Recombinant mammalian interleukin-8 receptor - used for screening
 PT Interleukin-8 binding antagonists, used to treat inflammation
 PS Claim 2; Fig 1; 71pp; English.
 CC Rabbit high affinity IL-8 receptor gene was isolated from rabbit
 CC peritoneal neutrophils and used as a source of poly(A)+ RNA, to
 CC produce a rabbit neutrophil cDNA library. 250,000 recombinant
 CC plaques were screened for those which hybridized to an antisense
 CC oligonucleotide (Q30015). This probe was designed based on the
 CC sequence derived from the second transmembrane domain of G-protein-
 CC coupled receptors. After tertiary screening, six plaques were
 CC isolated. The insert of one of these plaques, termed F3R was of 2.5
 CC kb in size. This insert was sequenced. The protein deduced from
 CC the F3R clone demonstrates that it belongs to the family of
 CC G-protein-coupled receptors. The deduced protein sequence
 CC indicates seven putative transmembrane segments.
 SQ Sequence 355 AA:

Query Match 2.8%; Score 10; DB 1; Length 355;
 Best Local Similarity 100.0%; Pred. No. 0.076;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 133 DRYLAIVHAT 142
 139 DRYLAIVHAT 148

RESULT 18

R28273 ID R28273 standard; Protein; 360 AA.

AC R28273; 04-APR-1993 (first entry)
 DE Sequence in a low affinity recombinant human interleukin-8
 DE (IL-8) receptor polypeptide in 4B.
 KW IL-8 receptor polypeptide; G-protein-coupled receptor.
 OS Homo sapiens.

PN W09218641-A.
 PD 29-OCT-1992.
 PF 10-APR-1992; U02977.
 PR 10-APR-1991; US-685101.
 PR 09-JUL-1991; US-726606.
 PR 09-DEC-1991; US-803842.
 PA (REPK) REPLIGEN CORP.
 PA (UYBO-) UNIV BOSTON.

PI Navarro J Thomas KM, Wilt DP;
 DR WPI: 92-362123/46.
 N-PSDB: Q30012.
 PT Recombinant mammalian Interleukin-8 receptor - used for screening
 PT Interleukin-8 binding antagonists, used to treat inflammation
 PS Disclosure: Fig 2: 71pp: English.
 CC Rabbit high affinity IL-8 receptor gene was isolated from rabbit
 CC peritoneal neutrophils and used as a source of poly(A)⁺ RNA, to
 CC produce a rabbit neutrophil cDNA library. 250,000 recombinant
 CC plaques were screened for those which hybridized to an antisense
 CC oligonucleotide (Q30015). This probe was designed based on the
 CC sequence derived from the second transmembrane domain of G-protein-
 CC coupled receptors. After tertiary screening, six plaques were
 CC isolated. The insert of one of these plaques, termed F3R was of 2.5
 CC kb in size. This insert was sequenced. The protein deduced from
 CC the F3R clone demonstrates that it belongs to the family of
 CC G-protein-coupled receptors. The deduced protein sequence
 CC indicates seven putative transmembrane segments. A human
 CC peripheral blood leukocyte lambda gtl cDNA library (5' stretch) was
 CC screened with a 652 bp EcoRI/BamHI fragment (including nucleotides
 CC -27 to 625) of the rabbit F3R clone. After tertiary screening
 CC several human clones which hybridized to the rabbit IL-8 probe
 CC were isolated. The insert of one such clone, termed 4AB was
 CC sequenced (Q30012).
 SO Sequence 360 AA;

Query Match 2.8%; Score 10; DB 1; Length 360;
 Best Local Similarity 100.0%; Pred. No. 0.077;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 133 DRYLAIVHAT 142
 |||||
 DB 143 DRYLAIVHAT 152

RESULT 19
 R33420
 ID R33420 standard; Protein: 355 AA.
 AC R33420:
 DT 26-JUL-1993 (first entry)
 DE Human IL-8 receptor from clone p2.
 KW Interleukin-8 receptor; probes; gene therapy; gro receptor;
 KW intracellular-calcium mobilising; ligand-binding; MIP-2 receptor.
 OS Homo sapiens.
 PN WO9306229-A.
 PD 01-APR-1993:
 PF 14-SEP-1992: U07641.
 PR 13-SEP-1991: US-759568.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
 PI Murphy PM;
 DR WPI: 93-117549/14.
 DR N-PSDB: Q38747.
 PT New interleukin-8 receptor aminoacid sequence - and corresp. cDNA
 PT expressed in Xenopus laevis oocytes or transfected host cells,
 PT for screening ligands of IL-8 receptor and gene therapy
 PS Claim 1: Fig 3: 39pp: English.
 CC cDNA libraries from 2 and 3.5 kb fractions of poly(A)⁺ RNA from HL60
 CC neutrophils sept. of a sucrose gradient were made in UniAmp. The
 CC libraries were screened with F3R oligonucleotide probe (from rabbit
 CC IL-8 receptor) and under low stringency with a p2 cDNA probe
 CC synthesised from random primers, to isolated the clone p2, encoding
 CC human IL-8 receptor protein. The presence or absence of the DNA
 CC coding IL-8R or related MIP-2 receptor may be detected using
 CC portions of the p2 clone as probes. p2 may also be used to screen
 CC for ligands of IL-8R and may also be used in gene therapy to treat a
 CC patient deficient in IL-8R. The IL-8R is a gro receptor and has
 CC intracellular calcium-mobilising and ligand-binding properties.
 SO Sequence 355 AA;

Query Match 2.8%; Score 10; DB 1; Length 355;
 Best Local Similarity 100.0%; Pred. No. 0.076;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 133 DRYLAIVHAT 142
 |||||
 DB 138 DRYLAIVHAT 147

RESULT 20
 ID R37903 standard; peptide: 18 AA.
 AC R37903:
 DT 05-NOV-1993 (first entry)
 DE Interleukin-8 receptor second cytoplasmic loop analogue.
 KW Inflammation; inflammatory response; G-protein-coupled receptor;
 KW phagocyte; chemotactic gradient; anti-inflammatory; IL-8-R;
 KW chemotactic receptor; formylpeptide receptor; immunosuppression.
 OS Synthetic.
 PN WO9311784-A.
 PD 24-JUN-1993:
 PF 11-DEC-1992: U10580.
 PR 13-DEC-1991: US-608453.
 PA (UYDU-) UNIV DUKE.
 PI Didsbury JR, Snyderman R, Uhlig RJ;
 DR WPI: 93-213817/26.
 PT New antiinflammatory or pro-inflammatory peptide(s) - inhibit
 PT chemo-attractant receptor-protein G interaction or receptor
 PT desensitisation, for treating arthritis, psoriasis,
 PT immunosuppression, etc.
 PS Claim 5; Page 59; 76pp: English.
 CC Analogue peptides based on the sequence of positively-charged
 CC regions in the second cytoplasmic loop of chemotactic receptors
 CC can inhibit the inflammatory response. Peptides were designed based
 CC on regions of homology on the second cytoplasmic loop of known
 CC receptors. R37903 is based on a sequence from the interleukin-8
 CC receptor. Peptides having this sequence or having a sequence of
 CC not more than 50 amino acids and comprising this 18mer are preferred
 CC anti-inflammatory compounds of the invention.
 CC See R37901-R37904.
 SO Sequence 18 AA;

Query Match 2.8%; Score 10; DB 1; Length 18;
 Best Local Similarity 100.0%; Pred. No. 0.0053;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 133 DRYLAIVHAT 142
 |||||
 DB 1 DRYLAIVHAT 10

RESULT 21
 ID R68811 standard; Protein: 350 AA.
 AC R68811:
 DT 18-JUL-1995 (first entry)
 DE Interleukin-8 receptor.
 KW Interleukin-8 receptor; IL-8 receptor; PF4R;
 KW platelet factor superfamily receptor; neutrophil; chemotactic;
 KW inflammation; inflammatory disease; arthritis; emphysema; cystic;
 KW fibrosis; colitis; bronchitis; meningitis; therapeutic.
 OS Homo sapiens.
 PN WO9428931-A.
 PD 22-DEC-1994:
 PF 07-JUN-1994: U06380.
 PR 11-JUN-1993: US-076093.
 PA (GETH) GENENTECH INC.
 PI Chuntarapai A, Hebert C, Kim KJ, Lee J;
 DR WPI: 95-036114/05.
 DR N-PSDB: Q80520.
 PT Treatment of inflammatory disorders - by administering an
 PT antibody capable of binding a platelet factor 4 superfamily
 PT receptor polypeptide
 PS Disclosure: Page 51-54; 83pp: English.

CC A cDNA library constructed from human neutrophil mRNA in pRKB was
 CC transfected into COS-7 cells, and the cells were screened with 125I-
 CC IL-8. The DNA sequence of isolated cDNA clone pRKB.118r1.1,
 CC encoding human IL-8 receptor, is given in 080520 and the predicted
 CC amino acid sequence in R68811. The receptor is used to raise
 CC antibodies that neutralize the activity of PF4AR, e.g. IL-8 receptor.
 SQ Sequence 350 AA;

Query Match 2.8%; Score 10; DB 1; Length 350;
 Best Local Similarity 100.0%; Pred. No. 0.075;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 133 DRYLAIVHAT 142
 |||||
 DB 134 DRYLAIVHAT 143

RESULT 22

R67671 standard: Protein; 367 AA.
 AC R67671:
 DI 17-AUG-1995 (first entry)
 DE Mouse opiod receptor-like receptor MOP2 protein.
 KM Mouse: kappa: delta: mu: opiod receptor: brain: primer: PCR: amplify:
 KW Transmembrane domain; somatostatin; receptor; human; expression vector;
 KW truncate: chimeric; assay; probe.
 OS Mus musculus.
 PN WO9428132-A.
 PD 08-DEC-1994.
 PF 20-MAY-1994; 005747.
 PR 20-MAY-1993; US-066296.
 PR 30-JUL-1993; US-100694.
 PR 05-NOV-1993; US-147592.
 PA (ARCH-) ARCH DEV CORP.
 PI Bell GI, Reisine T, Yasuda K;
 DR WPI: 95-022804/03.
 DR N-PSDB: 075928.

PT Polynucleotides and peptides derived from opiod receptor
 PT polypeptides - for use in therapeutic compositions and in
 PT screening assays for useful drug substances.
 PS Claim 27; Page 225-229; 300pp; English.
 CC The amino acid sequence of the novel mouse opiod receptor-like receptor
 CC MOP2. MOP2 is a mouse receptor with pharmacological properties which
 CC are dissimilar to the properties of classic opiod receptors such kappa,
 CC delta, mu or sigma. It has been found that drugs of high abuse potential
 CC or analgesic potency bind selectively to this receptor. This suggests
 CC that this receptor could be important in the development of drugs to
 CC treat addiction. Other opiod receptors isolated and produced such as the
 CC novel mouse kappa and delta opiod receptors (Q75926-7) are useful for
 CC the development of novel assays designed to select or improve substances,
 CC capable of interacting with the opiod receptor proteins, for use in
 CC diagnosis, drug design and therapeutic applications.
 SQ Sequence 367 AA;

Query Match 2.8%; Score 10; DB 1; Length 367;
 Best Local Similarity 100.0%; Pred. No. 0.078;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 296 CLNPILYAFL 305
 |||||
 DB 310 CLNPILYAFL 319

RESULT 23

R71968 standard: Protein; 367 AA.
 AC R71968:
 DI 20-OCT-1995 (first entry)
 DE Rat opiod receptor.
 KW Opiod receptor; gene therapy; diagnostic.
 OS Rattus sp.

PN WO9507983-A.
 PD 23-MAR-1995.
 PF 13-SEP-1994; U10358.
 PR 13-SEP-1993; US-120601.
 PA (INDV) UNIV INDIANA FOUND.

PI Yu L;
 DR WPI: 95-131351/17.
 DR N-PSDB: 089233.
 PT New nucleic acid encoding new human mu opiod receptor - and
 PT related vectors, transformed cells, antibodies etc., useful in
 PT diagnosis, treatment and drug screening.
 PS Example 9; Page 218-222; 266pp; English.
 CC The cDNA given in 089233 was isolated from a rat brain library by
 CC low stringency hybridization with rat mu opiod receptor cDNA
 CC (089222). The clone encoded a 367-amino acid protein (R71968)
 CC that showed high homology with mu, kappa and delta opiod receptors
 CC but lacked affinity for their ligands, suggesting it to be
 CC a novel member of the opiod receptor family.
 SQ Sequence 367 AA;

Query Match 2.8%; Score 10; DB 1; Length 367;
 Best Local Similarity 100.0%; Pred. No. 0.078;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 296 CLNPILYAFL 305
 |||||
 DB 310 CLNPILYAFL 319

RESULT 24

R74298 standard: Protein; 367 AA.
 AC R74298:
 DI 03-NOV-1995 (first entry)
 DE Mouse kappa-3 opiod receptor.
 KW kappa-3 opiod receptor; analgesia.
 OS Mus sp.
 PN WO9512616-A.
 PD 11-MAY-1995.
 PF 03-NOV-1994; U12728.
 PR 05-NOV-1993; US-147949.
 PA (SLOK) SLOAN KETTERING INST CANCER RES.
 PI Pan Y, Pasternak GW;
 DR WPI: 95-193814/25.
 DR N-PSDB: 090096.
 PT Nucleic acid molecule(s) encoding a kappa-3 opiod receptor, and
 PT antibody against the receptor - used to detect the receptor, and to
 PT image cell membrane-bound receptor in a subject
 PS Disclosure; Fig.1; 68pp; English.
 CC Degenerate primers based on conserved sequences of the mouse delta
 CC opiod receptor were used in PCR to amplify mouse kappa-3 opiod
 CC receptor cDNA from a mouse brain lambda ZAP cDNA library. The
 CC encoded protein was expressed in E. coli and COS-7 cells.
 SQ Sequence 367 AA;

Query Match 2.8%; Score 10; DB 1; Length 367;
 Best Local Similarity 100.0%; Pred. No. 0.078;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 296 CLNPILYAFL 305
 |||||
 DB 310 CLNPILYAFL 319

RESULT 25

R76638 standard: Protein; 367 AA.
 AC R76638:
 DI 22-DEC-1995 (first entry)
 DE Rat opiorph receptor OR7.
 KW Opiorph receptor; opiod.

OS Rattus sp.
 PN W09519986-A1.
 PD 27-JUL-1995.
 PF 20-JAN-1995; U00939.
 PR 21-JAN-1994; US-185360.
 PA (AMCY) AMERICAN CYANAMID CO.
 PI Eppler CM, Hulmes JD, Ozenberger BA;
 DR WPI: 95-269412/35.
 DR N-PSDB: Q92972.
 PT New isolated DNA encoding an opiorph receptor - used to develop
 PT prods. for identifying opiorph agonists and antagonists and for
 PT detection and manipulation
 PS Disclosure: Fig.1: 35pp; English.
 CC Primers based on rodent opiorph receptors were used to amplify rat
 CC genomic DNA. Products were re-amplified and subcloned into PCR-II
 CC vector and amplified in E. coli. Plasmid DNAs were isolated, and
 CC an unspliced sequence was obtd. (Q92972) encoding the transmembrane
 CC domain opiorph receptor protein OR7.
 SO Sequence 367 AA:

Query Match 2.8%; Score 10; DB 1; Length 367;
 Best Local Similarity 100.0%; Pred. No. 0.078;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 296 CLNPILYAF 305
 |||||
 Db 310 CLNPILYAF 319

RESULT 26
 R70123
 ID R70123 standard; Protein; 1060 AA.
 AC R70123;
 DT 14-FEB-1996 (first entry)
 DE IL8-R type 1-GBP 130 fusion protein.
 KW Hybrid peptide; malaria parasite; Plasmodium falciparum; fusion protein;
 KW red blood cell; cytokine receptor; glycophorin binding peptide 130;
 KW GBP 130; GBPH; glycophorin binding peptide homologue; glycophorin A.
 OS Chimeric Homo sapiens.
 OS Chimeric Plasmodium falciparum.
 PN W09506737-A.
 PD 09-MAR-1995.
 PF 01-SEP-1994; G01900.
 PR 03-SEP-1993; GB-018350.
 PR 23-AUG-1994; GB-017021.
 PA (PREN/) PRENDERGAST K F.
 PI Prendergast KF;
 DR WPI: 95-115452/15.
 PT New hybrid peptide(s) for binding cytokine(s) - comprising a
 PT malaria parasite peptide capable of binding a red blood cell and
 PT a receptor peptide.
 PS Example A: Page 79-80; 93pp; English.
 CC Hybrid peptides for binding cytokines, comprising a malaria parasite
 CC (Plasmodium falciparum) peptide (capable of binding to a red blood
 CC cell (RBC)) and a receptor peptide are claimed. R70103-25 are examples
 CC of these hybrid peptides. R70123 is a fusion of Interleukin 8 receptor
 CC type 1 and glycophorin binding protein (GBP) 130. The use of cytokine
 CC receptors not normally found on RBCs means that the cytokine can bind
 CC harmlessly to the RBC without deleterious effect. The RBC protects the
 CC hybrid peptides from excretion from the kidney, and due to steric
 CC hindrance prevents the cytokines binding to a receptor in another cell.
 CC GBP 130 or GBPH (GBP homologue) are the prefid. malaria parasite peptides
 CC used, others include EBA 175 (175 kda erythrocyte binding antigen),
 CC PMMSA (pre major merozoite surface antigen) and the Duffy binding
 CC receptor molecule (eg. exhibited by Plasmodium vivax). These peptides
 CC bind to pref. glycophorin A, B and C, sialo glycoproteins, found on
 CC the surface of RBCs. The hybrid peptides are thus used to lower
 CC the levels of free cytokines in the circulation to reduce pathological
 CC damage.
 SO Sequence 1060 AA:

Query Match 2.8%; Score 10; DB 1; Length 1060;
 Best Local Similarity 100.0%; Pred. No. 0.2;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 133 DRYLAIVHAT 142
 |||||
 Db 134 DRYLAIVHAT 143

RESULT 27
 R70124
 ID R70124 standard; Protein; 1064 AA.
 AC R70124;
 DT 14-FEB-1996 (first entry)
 DE IL8-R type 2-GBP 130 fusion protein.
 KW Hybrid peptide; malaria parasite; Plasmodium falciparum; fusion protein;
 KW red blood cell; cytokine receptor; glycophorin binding peptide 130;
 KW GBP 130; GBPH; glycophorin binding peptide homologue; glycophorin A.
 OS Chimeric Homo sapiens.
 OS Chimeric Plasmodium falciparum.
 PN W09506737-A.
 PD 09-MAR-1995.
 PF 01-SEP-1994; G01900.
 PR 03-SEP-1993; GB-018350.
 PR 23-AUG-1994; GB-017021.
 PA (PREN/) PRENDERGAST K F.
 PI Prendergast KF;
 DR WPI: 95-115452/15.
 PT New hybrid peptide(s) for binding cytokine(s) - comprising a
 PT malaria parasite peptide capable of binding a red blood cell and
 PT a receptor peptide.
 PS Example A: Page 80-82; 93pp; English.
 CC Hybrid peptides for binding cytokines, comprising a malaria parasite
 CC (Plasmodium falciparum) peptide (capable of binding to a red blood
 CC cell (RBC)) and a receptor peptide are claimed. R70103-25 are examples
 CC of these hybrid peptides. R70124 is a fusion of Interleukin 8 receptor
 CC type 2 and glycophorin binding protein (GBP) 130. The use of cytokine
 CC receptors not normally found on RBCs means that the cytokine can bind
 CC harmlessly to the RBC without deleterious effect. The RBC protects the
 CC hybrid peptides from excretion from the kidney, and due to steric
 CC hindrance prevents the cytokines binding to a receptor in another cell.
 CC GBP 130 or GBPH (GBP homologue) are the prefid. malaria parasite peptides
 CC used, others include EBA 175 (175 kda erythrocyte binding antigen),
 CC PMMSA (pre major merozoite surface antigen) and the Duffy binding
 CC receptor molecule (eg. exhibited by Plasmodium vivax). These peptides
 CC bind to pref. glycophorin A, B and C, sialo glycoproteins, found on
 CC the surface of RBCs. The hybrid peptides are thus used to lower
 CC the levels of free cytokines in the circulation to reduce pathological
 CC damage.
 SO Sequence 1064 AA:

Query Match 2.8%; Score 10; DB 1; Length 1064;
 Best Local Similarity 100.0%; Pred. No. 0.2;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 133 DRYLAIVHAT 142
 |||||
 Db 137 DRYLAIVHAT 146

RESULT 28
 R80756
 ID R80756 standard; Protein; 350 AA.
 AC R80756;
 DT 26-MAR-1996 (first entry)
 DE Interleukin 8 receptor A partial sequence.
 KW Interleukin; IL-8; inflammation; psoriasis; dermatitis;
 KW rheumatoid arthritis; inflammatory bowel disease;
 KW chronic lung inflammation; treatment; antibody;
 KW affinity purification; detection.
 OS Homo sapiens.
 PN U55440021-A.

PD 08-AUG-1995: 677211.
 PF 29-MAR-1991: US-677211.
 PR 29-MAR-1991: US-677211.
 PR 25-FEB-1994: US-202056.
 PA (CHUN/) CHUNTHARAPAI A.
 PA (HEBE/) HERBERT C.
 PA (KIMK/) KIM K J.
 PA (LEEJ/) LEE J.
 PI Chuntharapai A, Herbert C, Kim KJ, Lee J;
 DR WPI: 95-283151/37.
 DR N-PSDB: Q99006.
 PT New antibodies against interleukin 8 type B receptor - used to treat or prevent inflammation, also for detecting receptor expression and purification.
 PT Example 2: Columns 41-44: 62pp: English.
 PS Antibodies directed against the interleukin-8 receptor B can be used to treat or prevent inflammation e.g. psoriasis, dermatitis, rheumatoid arthritis and particularly inflammatory bowel disease and chronic lung inflammation. When immobilised, these antibodies may be used to detect interleukin-8 receptor B expression in cells and tissues and for affinity purification of interleukin-8 receptor B from cells. A sequence encoding amino acids 23-314 of this fragment of the high affinity interleukin-8 receptor, was used to probe lambda gt10 cDNA libraries for the human interleukin-8 type B receptor.
 SQ Sequence 350 AA:

Query Match 2.8%; Score 10; DB 1; Length 350;
 Best Local Similarity 100.0%; Pred. No. 0.075;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 133 DRYLAIVHAT 142
 |||||
 DB 134 DRYLAIVHAT 143

RESULT 29
 R80758
 ID R80758 standard; Protein; 360 AA.
 AC R80758:
 DI 26-MAR-1996 (first entry)
 DE Interleukin 8 receptor B.
 KW Interleukin: IL-8; inflammation; psoriasis; dermatitis;
 KW rheumatoid arthritis; inflammatory bowel disease;
 KW chronic lung inflammation; treatment; antibody;
 KW affinity purification; detection.
 OS Homo sapiens.
 PN US5440021-A.
 PD 08-AUG-1995.
 PF 29-MAR-1991: 677211.
 PR 29-MAR-1991: US-677211.
 PR 25-FEB-1994: US-202056.
 PA (CHUN/) CHUNTHARAPAI A.
 PA (HEBE/) HERBERT C.
 PA (KIMK/) KIM K J.
 PA (LEEJ/) LEE J.
 PI Chuntharapai A, Herbert C, Kim KJ, Lee J;
 DR WPI: 95-283151/37.
 DR N-PSDB: Q99008.
 PT New antibodies against interleukin 8 type B receptor - used to treat or prevent inflammation, also for detecting receptor expression and purification.
 PS Disclosure: Columns 53-56: 62pp: English.
 CC Antibodies directed against the interleukin-8 receptor B can be used to treat or prevent inflammation e.g. psoriasis, dermatitis, rheumatoid arthritis and particularly inflammatory bowel disease and chronic lung inflammation. When immobilised, these antibodies may be used to detect interleukin-8 receptor B expression in cells and tissues and for affinity purification of interleukin-8 receptor B from cells.
 SQ Sequence 360 AA:

Query Match 2.8%; Score 10; DB 1; Length 360;
 Best Local Similarity 100.0%; Pred. No. 0.077;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 133 DRYLAIVHAT 142
 |||||
 DB 143 DRYLAIVHAT 152

RESULT 30
 R80953
 ID R80953 standard; Protein; 360 AA.
 AC R80953:
 DI 24-APR-1996 (first entry)
 DE Recombinant high affinity interleukin-8 receptor subtype B.
 KW IL-8A; IL-8B; receptor; monoclonal antibody; inflammatory disorder;
 KW anaphylaxis; systemic lupus erythematosus; rheumatoid arthritis;
 KW systemic necrotizing vasculitis; psoriasis; asthma; allergy; ARDS;
 KW adult respiratory distress syndrome; neutrophil detection.
 OS Homo sapiens.
 PN W09525126-A1.
 PD 21-SEP-1995.
 PF 09-MAR-1995: U03032.
 PR 15-MAR-1994: US-210250.
 PR 02-MAY-1994: US-237937.
 PA (REPK) REPLIGEN CORP.
 PA (UYBO-) UNIV BOSTON.
 PI Greenfield EA, Larosa GJ, Navarro J, Thomas KM;
 PI Wilt DT
 DR WPI: 95-336945/43.
 DR N-PSDB: Q99952.
 PT Monoclonal antibody against recombinant IL-8 receptor polypeptide - useful for treating inflammatory disorders, for detecting neutrophil(s) and for isolating IL-8 receptor from liq.mixt.
 PS Claim 5: Fig 4A-B: 74pp: English.
 CC Monoclonal antibodies were raised against recombinant interleukin-8 (IL-8) receptor subtypes A and B from both human and rabbit sources (R80950-53 encoded by Q99949-52). The A subtype receptor (IL-8A) is a high affinity receptor and the B subtype receptor (IL-8B) is a low affinity receptor. The monoclonal antibody (mAb) pref. binds to the IL-8 binding domain thus blocking its activation. The mAbs are useful for treating inflammatory disorders (see key words) and for detecting the presence of neutrophils in a biological sample. The mAbs are also useful in the isolation of IL-8 receptors from a mixture.
 SQ Sequence 360 AA:

Query Match 2.8%; Score 10; DB 1; Length 360;
 Best Local Similarity 100.0%; Pred. No. 0.077;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 133 DRYLAIVHAT 142
 |||||
 DB 143 DRYLAIVHAT 152

RESULT 31
 R80950
 ID R80950 standard; Protein; 355 AA.
 AC R80950:
 DI 24-APR-1996 (first entry)
 DE Recombinant high affinity interleukin-8 receptor subtype A.
 KW IL-8A; IL-8B; receptor; monoclonal antibody; inflammatory disorder;
 KW anaphylaxis; systemic lupus erythematosus; rheumatoid arthritis;
 KW systemic necrotizing vasculitis; psoriasis; asthma; allergy; ARDS;
 KW adult respiratory distress syndrome; neutrophil detection.
 OS Oryzolaegus cuniculus.
 PN W09525126-A1.
 PD 21-SEP-1995.
 PF 09-MAR-1995: U03032.
 PR 15-MAR-1994: US-210250.
 PR 02-MAY-1994: US-237937.
 PA (REPK) REPLIGEN CORP.

PA (UYBO-) UNIV BOSTON.
 PI Greenfield EA, Larosa GJ, Navarro J, Thomas KM;
 PI Wilt DT;
 DR WPI: 95-336945/43.
 DR N-PSDB: Q99949.
 PT Monoclonal antibody against recombinant IL-8 receptor polypeptide -
 useful for treating inflammatory disorders, for detecting
 PI neutrophil(s) and for isolating IL-8 receptor from liq.mixt.
 PS Claim 2: Fig 3A-B: 74pp; English.
 CC Monoclonal antibodies were raised against recombinant interleukin-8
 CC (IL-8) receptor subtypes A and B from both human and rabbit sources
 CC (R80950-53 encoded by Q99949-52). The A subtype receptor (IL-8RA) is
 CC a high affinity receptor and the B subtype receptor (IL-8RB) is a
 CC low affinity receptor. The monoclonal antibody (mAb) pref. binds to
 CC the IL-8 binding domain thus blocking its activation. The mAbs are
 CC useful for treating inflammatory disorders (see key words) and for
 CC detecting the presence of neutrophils in a biological sample. The
 CC mAbs are also useful in the isolation of IL-8 receptors from a mixture.
 SQ Sequence 355 AA;

Query Match 2.8%; Score 10; DB 1; Length 355;
 Best Local Similarity 100.0%; Pred. No. 0.076;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 133 DRYLAIWAT 142
 Db 139 DRYLAIWAT 148

RESULT 32
 ID R80951
 AC R80951: standard; Protein: 350 AA.
 DT 24-APR-1996 (first entry)
 DE Recombinant high affinity interleukin-8 receptor subtype A.
 KW IL-8A; IL-8B; receptor; monoclonal antibody; inflammatory disorder;
 KW anaphylaxis; systemic lupus erythematosus; rheumatoid arthritis;
 KW systemic necrotizing vasculitis; psoriasis; asthma; allergy; ARDS;
 KW adult respiratory distress syndrome; neutrophil detection.
 OS Homo sapiens.
 PN W09525126-A1.
 PD 21-SEP-1995.
 PR 09-MAR-1995: U03032.
 PR 15-MAR-1994: US-210250.
 PR 02-MAY-1994: US-237937.
 PA (REPK) REPLIGEN CORP.
 PA (UYBO-) UNIV BOSTON.
 PI Greenfield EA, Larosa GJ, Navarro J, Thomas KM;
 PI Wilt DT;
 DR WPI: 95-336945/43.
 DR N-PSDB: Q99950.
 PT Monoclonal antibody against recombinant IL-8 receptor polypeptide -
 useful for treating inflammatory disorders, for detecting
 PI neutrophil(s) and for isolating IL-8 receptor from liq.mixt.
 PS Disclosure: Fig 2A-B: 74pp; English.
 CC Monoclonal antibodies were raised against recombinant interleukin-8
 CC (IL-8) receptor subtypes A and B from both human and rabbit sources
 CC (R80950-53 encoded by Q99949-52). The A subtype receptor (IL-8RA) is
 CC a high affinity receptor and the B subtype receptor (IL-8RB) is a
 CC low affinity receptor. The monoclonal antibody (mAb) pref. binds to
 CC the IL-8 binding domain thus blocking its activation. The mAbs are
 CC useful for treating inflammatory disorders (see key words) and for
 CC detecting the presence of neutrophils in a biological sample. The
 CC mAbs are also useful in the isolation of IL-8 receptors from a mixture.
 SQ Sequence 350 AA;

Query Match 2.8%; Score 10; DB 1; Length 350;
 Best Local Similarity 100.0%; Pred. No. 0.075;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 133 DRYLAIWAT 142

Db 134 DRYLAIWAT 143

RESULT 33
 ID R80952
 AC R80952: standard; Protein: 358 AA.
 DT 24-APR-1996 (first entry)
 DE Recombinant high affinity interleukin-8 receptor subtype B.
 KW IL-8A; IL-8B; receptor; monoclonal antibody; inflammatory disorder;
 KW anaphylaxis; systemic lupus erythematosus; rheumatoid arthritis;
 KW systemic necrotizing vasculitis; psoriasis; asthma; allergy; ARDS;
 KW adult respiratory distress syndrome; neutrophil detection.
 OS Oryzolaus cuniculus.
 PN W09525126-A1.
 PD 21-SEP-1995.
 PR 09-MAR-1995: U03032.
 PR 15-MAR-1994: US-210250.
 PR 02-MAY-1994: US-237937.
 PA (REPK) REPLIGEN CORP.
 PA (UYBO-) UNIV BOSTON.
 PI Greenfield EA, Larosa GJ, Navarro J, Thomas KM;
 PI Wilt DT;
 DR WPI: 95-336945/43.
 DR N-PSDB: Q99951.
 PT Monoclonal antibody against recombinant IL-8 receptor polypeptide -
 useful for treating inflammatory disorders, for detecting
 PI neutrophil(s) and for isolating IL-8 receptor from liq.mixt.
 PS Claim 6: Fig 3A-B: 74pp; English.
 CC Monoclonal antibodies were raised against recombinant interleukin-8
 CC (IL-8) receptor subtypes A and B from both human and rabbit sources
 CC (R80950-53 encoded by Q99949-52). The A subtype receptor (IL-8RA) is
 CC a high affinity receptor and the B subtype receptor (IL-8RB) is a
 CC low affinity receptor. The monoclonal antibody (mAb) pref. binds to
 CC the IL-8 binding domain thus blocking its activation. The mAbs are
 CC useful for treating inflammatory disorders (see key words) and for
 CC detecting the presence of neutrophils in a biological sample. The
 CC mAbs are also useful in the isolation of IL-8 receptors from a mixture.
 SQ Sequence 358 AA;

Query Match 2.8%; Score 10; DB 1; Length 358;
 Best Local Similarity 100.0%; Pred. No. 0.077;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 133 DRYLAIWAT 142
 Db 141 DRYLAIWAT 150

RESULT 34
 ID W25217
 AC W25217: standard; peptide: 367 AA.
 DT 12-JAN-1998 (first entry)
 DE Rat orphanin FQ receptor encoded by cDNA clone LC132.
 KW Orphanin FQ receptor; binding; locomotor disease; diagnosis;
 KW treatment; opioid inhibitor; opiate induced hypothermia; drug design;
 KW morphine induced analgesia; methadone specific opioid receptor; ss.
 OS Rattus sp.
 FH Key
 FT Misc_difference 21 Location/Qualifiers
 FT /note= "potential N-linked glycosylation site"
 FT Misc_difference 26 /note= "potential N-linked glycosylation site"
 FT /note= "potential N-linked glycosylation site"
 FT Misc_difference 36 /note= "potential N-linked glycosylation site"
 FT /note= "potential N-linked glycosylation site"
 FT Domain 45..75 /note= "transmembrane domain I"
 FT Domain 81..106 /note= "transmembrane domain II"
 FT /note= "transmembrane domain II"
 FT Domain 122..143

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 8, 1999, 02:59:26 ; Search time 11.06 Seconds

(without alignments)
1275.140 Million cell updates/sec

Title: US-09-104-063-4

Sequence: 1 MEGISITYSDNYTEEMGSD.....KRGHSSVSTESSESSFHSS 352

Scoring table: OLIGO

Searched: 122810 segs, 40065486 residues

Database: PIR_60:*

1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	352	100.0	352	2	A45747	neuropeptide Y/pep
2	114	32.4	352	2	G00048	fusin (LESTRA) - c
3	74	21.0	353	2	S28787	neuropeptide Y/pep
4	10	2.8	350	2	A39445	interleukin-8 rece
5	10	2.8	360	2	A53611	interleukin-8 rece
6	10	2.8	370	2	S43087	orphan opioid rece
7	10	2.8	358	2	A53752	interleukin-8 rece
8	10	2.8	355	2	JQ1231	interleukin-8 rece
9	10	2.8	359	2	A48921	interleukin-8 rece
10	10	2.8	367	2	I49022	K3 opiate receptor
11	10	2.8	367	2	JC2421	opioid receptor ho
12	10	2.8	367	2	I56520	G protein-coupled
13	10	2.8	356	2	S42096	interleukin-8 rece

ALIGNMENTS

RESULT 1
A45747
neuropeptide Y/peptide YY receptor Y3 - human
N:Alternate names: fusin; HM99; leukocyte-derived seven-transmembrane receptor LESTR; re
C:Species: Homo sapiens (man)
C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 17-Mar-1999
C:Accession: A45747; A53103; I53006; I59444; I69203; S33761
R:Feierspiel, B.; Melhado, I.G.; Duncan, A.M.V.; Delaney, A.; Schappert, K.; Clark-Lewis
Genomics 16, 707-712, 1993
A:Title: Molecular cloning of the cDNA and chromosomal localization of the gene for a pu
A:Reference number: A45747; MUID:93315154
A:Accession: A45747
A:Molecule type: mRNA
A:Residues: 1-352 <FEU>
A:Cross-references: GB:M99293; NID:9292516; PID:9292517
R:Loetscher, M.; Geisler, T.; O'Reilly, T.; Zechlin, R.; Baggiolini, M.; Moser, B.
J. Biol. Chem. 269, 232-237, 1994
A:Title: Cloning of a human seven-transmembrane domain receptor, LESTR, that is highly e
A:Reference number: A53103; MUID:94103215
A:Accession: A53103
A:Molecule type: mRNA

A:Residues: 1-352 <LOE>
A:Cross-references: EMBL:X71635; NID:9297099; PID:9297100
R:Herzog, H.; Hort, J.O.; Shine, J.; Seibler, L.A.
DNA Cell Biol. 12, 465-471, 1993
A:Title: Molecular cloning, characterization, and localization of the human homolog t
A:Reference number: I53006; MUID:93319629
A:Accession: I53006
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-352 <HER>
A:Cross-references: GB:L06797; NID:9414929; PID:9414928
R:Jazin, E.E.; Yoo, H.; Blomqvist, A.G.; Yee, F.; Weng, G.; Walker, M.W.; Salton, J.;
Regul. Pept. 47, 247-258, 1993
A:Title: A proposed bovine neuropeptide Y (NPY) receptor cDNA clone, or its human hom
A:Reference number: I59444; MUID:9405283
A:Accession: I59444
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-352 <RE2>
A:Cross-references: GB:L01639; NID:9189313; PID:9189314
R:Nomura, H.; Nielsen, B.W.; Matsushima, K.
Int. Immunol. 5, 1239-1249, 1993
A:Title: Molecular cloning of cDNAs encoding a LD78 receptor and putative leukocyte c
A:Reference number: I54751; MUID:94092629
A:Accession: I69203
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-352 <RES>
A:Cross-references: GB:DI0924; NID:9219868; PID:9219869
C:Genetics:
A:Gene: GDB:NPY3R; NPY3
A:Cross-references: GDB:230002; OMIM:162643
A:Map position: 2q21-2q21
C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 100.0%; Score 352; DB 2; Length 352;
Best local similarity 100.0%; Pred. No. 0;
Matches 352; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEGISITYSDNYTEEMGSDYDSMKPCFRENNANFKFLPTIYSIFLTGIVNGLYI 60
|||||
Db 1 MEGISITYSDNYTEEMGSDYDSMKPCFRENNANFKFLPTIYSIFLTGIVNGLYI 60
QY 61 LVWGYCKLRSMYDKYRLSLVADLLFVTLPEMAADAVANNYEGNFKAVHYITVNL 120
|||||
Db 61 LVWGYCKLRSMYDKYRLSLVADLLFVTLPEMAADAVANNYEGNFKAVHYITVNL 120
QY 121 YSSVLLIAFISLDRLAIYHATNSORPRKRLAEKVYVGVWIPALLITIPDFIFANVSEA 180
|||||
Db 121 YSSVLLIAFISLDRLAIYHATNSORPRKRLAEKVYVGVWIPALLITIPDFIFANVSEA 180
QY 181 DDRTYCDREFPNDLWVYVQFOHIMWGLLPGLVILSCYIIISKISHSGQKRAKLT 240
|||||
Db 181 DDRTYCDREFPNDLWVYVQFOHIMWGLLPGLVILSCYIIISKISHSGQKRAKLT 240
QY 241 TYLLIAFAFACMLPYIYIGISIDFILLEIKOGCEPENVHWMISTEALAFHCCLNPI 300
|||||
Db 241 TYLLIAFAFACMLPYIYIGISIDFILLEIKOGCEPENVHWMISTEALAFHCCLNPI 300
QY 301 LYAFUGAKRTSAOHALTSVSRGSSILKISKGRGSHSVSTESSESSFHSS 352
|||||
Db 301 LYAFUGAKRTSAOHALTSVSRGSSILKISKGRGSHSVSTESSESSFHSS 352

RESULT 2
G00048
Fusin (LESTRA) - crab-eating macaque
C:Species: Macaca fascicularis (crab-eating macaque)
C:Date: 11-Apr-1997 #sequence_revision 11-Apr-1997 #text_change 11-Apr-1997
C:Accession: G00048
R:Tatsumi, M.
submitted to GenBank, July 1996

A:Reference number: H00048
 A:Accession: G00048
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-352 <TAR>
 A:Cross-references: GB:D86579; NID:g1468948; PID:g1468949

Query Match 32.4%; Score 114; DB 2; Length 352;
 Best Local Similarity 100.0%; Pred. No. 1.8e-107;
 Matches 114; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 177 VSEADRRICDFYPNDLVWVVFQHIWGLILPGIVILSCYCIISKLSHSGHOKRK 236
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 177 VSEADRRICDFYPNDLVWVVFQHIWGLILPGIVILSCYCIISKLSHSGHOKRK 236
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Oy 237 ALKTVIILAFACMLPYIGISIDSFLLEIKOGCEFEVTHKWSITLAL 290
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 237 ALKTVIILAFACMLPYIGISIDSFLLEIKOGCEFEVTHKWSITLAL 290
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 3
 S28787
 neuropeptide Y/peptide YY receptor Y3 - bovine
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 06-Dec-1996
 C:Accession: S28787
 R:Rimland, J.; Xin, W.; Sweetnam, P.; Saijoh, K.; Nestler, E.J.; Duman, R.S.
 Mol. Pharmacol. 40, 869-875, 1991
 A:Title: Sequence and expression of a neuropeptide Y receptor cDNA.
 A:Reference number: S28787; MUID:92100053
 A:Accession: S28787
 A:Molecule type: mRNA
 A:Residues: 1-353 <RIM>
 A:Cross-references: EMBL:M86739
 C:Keywords: appetite; G protein-coupled receptor; transmembrane protein

Query Match 21.0%; Score 74; DB 2; Length 353;
 Best Local Similarity 100.0%; Pred. No. 5.6e-67;
 Matches 74; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 279 TVHKKISITFEALAFHCCNPLIYAFILGAKFKTSQAHATSVSRGSSLTILSKRGKGS 338
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 280 TVHKKISITFEALAFHCCNPLIYAFILGAKFKTSQAHATSVSRGSSLTILSKRGKGS 339
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Oy 339 SVSTESSESSFFHSS 352
 ||||||||||||||||
 Db 340 SVSTESSESSFFHSS 353
 ||||||||||||||||

RESULT 4
 A39445
 Interleukin-8 receptor type A - human
 N:Alternate names: Interleukin-8 receptor, high-affinity
 C:Species: Homo sapiens (man)
 C:Date: 22-Jan-1993 #sequence_revision 12-Apr-1996 #text_change 24-Sep-1998
 C:Accession: I37449; I38710; I38711; A39445
 R:MollerEAU, C.; Muscatelli, F.; Mattei, M.G.; Vassart, G.; Parmentier, M.
 Genomics 16, 248-251, 1993
 A:Title: The high-affinity Interleukin 8 receptor gene (IL8RA) maps to the 2q33-q36 reg1
 A:Reference number: I37449; MUID:93252387
 A:Accession: I37449
 A:Molecule type: DNA
 A:Residues: 1-350 <RES>
 A:Cross-references: EMBL:X65858; NID:g312046; PID:g312047
 R:Ahuja, S.K.; Shetty, A.; Tiffany, H.L.; Murphy, P.M.
 J. Biol. Chem. 269, 26381-26389, 1994
 A:Title: Comparison of the genomic organization and promoter function for human interleukin-8 receptor type A.
 A:Reference number: I37898; MUID:95014476
 A:Accession: I38710
 A:Molecule type: DNA
 A:Residues: 1-350 <RE2>

A:Cross-references: EMBL:U11870; NID:g511804; PID:g511805
 A:Accession: I38711
 A:Molecule type: mRNA
 A:Residues: 1-16 <RE3>
 A:Cross-references: EMBL:U11871; NID:g511806; PID:g733002
 R:Holmes, W.E.; Lee, J.; Kuang, W.J.; Rice, G.C.; Wood, W.I.
 Science 253, 1278-1280, 1991
 A:Title: Structure and functional expression of a human interleukin-8 receptor.
 A:Reference number: A39445; MUID:91368199
 A:Accession: A39445
 A:Molecule type: mRNA
 A:Residues: 1-275, 'T', 277-350 <HOL>
 A:Cross-references: GB:M68932; NID:g186369; PID:g186370
 C:Genetics:
 A:Gene: GDB:IL8RA
 A:Cross-references: GDB:I35039; OMIM:146929
 A:Map position: 2q35-2q35
 C:Keywords: G protein-coupled receptor; glycoprotein; membrane protein

Query Match 2.8%; Score 10; DB 2; Length 350;
 Best Local Similarity 100.0%; Pred. No. 0.035;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 133 DRYLAIVHAT 142
 |||||||||
 Db 134 DRYLAIVHAT 143
 |||||||||

RESULT 5
 A53611
 Interleukin-8 receptor type B - human
 C:Species: Homo sapiens (man)
 C:Date: 07-Oct-1994 #sequence_revision 12-Apr-1996 #text_change 29-Aug-1997
 C:Accession: I37898; I38712; A53611; A39446
 R:Ahuja, S.K.; Shetty, A.; Tiffany, H.L.; Murphy, P.M.
 J. Biol. Chem. 269, 26381-26389, 1994
 A:Title: Comparison of the genomic organization and promoter function for human interleukin-8 receptor type B.
 A:Reference number: I37898; MUID:95014476
 A:Accession: I37898
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-360 <RES>
 A:Cross-references: EMBL:U11869; NID:g511801; PID:g511803
 A:Accession: I38712
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-15 <RE2>
 A:Cross-references: EMBL:U11872; NID:g511808; PID:g511809; EMBL:U11873; NID:g511810;
 6; PID:g511817; EMBL:U11877; NID:g511818; PID:g511819; EMBL:U11878; NID:g511820; PID:
 R:Sprengr, H.; Lloyd, A.R.; Lautens, L.L.; Bonner, T.I.; Kelvin, D.J.
 J. Biol. Chem. 269, 11065-11072, 1994
 A:Title: Structure, genomic organization, and expression of the human interleukin-8 r

A:Reference number: A53611; MUID:94209273
 A:Accession: A53611
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 6-360 <SPR>
 A:Cross-references: GB:M99412; GB:LI9593
 R:Murphy, P.M.; Tiffany, H.L.
 Science 253, 1280-1283, 1991
 A:Title: Cloning of complementary DNA encoding a functional human interleukin-8 recep
 A:Reference number: A39446; MUID:91368200
 A:Accession: A39446
 A:Status: preliminary; nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 6-360 <MUR>
 A:Cross-references: GB:M73969
 C:Comment: This receptor, unlike IL8RA, binds several peptides besides interleukin-8,
 C:Genetics:
 A:Gene: GDB:IL8RB; IL8RA
 A:Cross-references: GDB:I27868; OMIM:146928
 A:Map position: 2q35-2q35

C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 2.8%; Score 10; DB 2; Length 360;
Best Local Similarity 100.0%; Pred. No. 0.036;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 133 DRYLAIVHAT 142
|||||
Db 143 DRYLAIVHAT 152

RESULT 6

Orphan opioid receptor ORL1 - human
C:Species: Homo sapiens (man)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 10-Sep-1997
C:Accession: S43087; G01892
R:Moilleau, C.; Parmentier, M.; Mailleux, P.; Butour, J.L.; Moisand, C.; Chalon, P.; Ca
FEBS Lett. 341, 33-38, 1994
A:Title: ORL1, a novel member of the opioid receptor family. Cloning, functional express
A:Reference number: S43087; MUID:94185768
A:Accession: S43087
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-370 <MOL>
A:Cross-references: EMBL:X77130; NID:g471316; PID:g471317
R:Lee P., H.
submitted to the EMBL Data Library, June 1995
A:Reference number: H00703
A:Accession: G01892
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-370 <LEF>
A:Cross-references: EMBL:U30185; NID:g1144296; PID:g1144297
C:Genetics:
A:Gene: GDB:OPRL1
A:Cross-references: GDB:345029
A:Map position: Bq11.2-Bq11.2
C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 2.8%; Score 10; DB 2; Length 370;
Best Local Similarity 100.0%; Pred. No. 0.037;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 296 CLNPITYAFL 305
|||||
Db 313 CLNPITYAFL 322

RESULT 7

Interleukin-8 receptor (clone 5B1a) - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 10-Sep-1997
C:Accession: A53752
R:Prado, G.N.; Thomas, K.M.; Suzuki, H.; LaRosa, G.J.; Wilkinson, N.; Folco, E.; Navaric
J. Biol. Chem. 269, 12391-12394, 1994
A:Title: Molecular characterization of a novel rabbit interleukin-8 receptor isotype.
A:Reference number: A53752; MUID:94230294
A:Accession: A53752
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-358 <PRA>
A:Cross-references: GB:L24445; NID:g437661; PID:g437662
C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 2.8%; Score 10; DB 2; Length 358;
Best Local Similarity 100.0%; Pred. No. 0.036;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 133 DRYLAIVHAT 142
|||||
Db 141 DRYLAIVHAT 150

RESULT 8

Interleukin-8 receptor - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 20-Mar-1998
C:Accession: J01231; A46483
R:Beckmann, M.P.; Munger, W.E.; Kozlosky, C.; Vandembos, T.; Price, V.; Lyman, S.; Ge
Biochem. Biophys. Res. Commun. 179, 784-789, 1991
A:Title: Molecular characterization of the Interleukin-8 receptor.
A:Reference number: J01231; MUID:91378994
A:Accession: J01231
A:Molecule type: DNA
A:Residues: 1-355 <BEC>
A:Cross-references: GB:M74240; NID:g165438; PID:g165439
R:Lee, J.; Kuang, W.J.; Rice, G.C.; Wood, W.I.
J. Immunol. 148, 1261-1264, 1992
A:Title: Characterization of complementary DNA clones encoding the rabbit IL-8 recept
A:Reference number: A46483; MUID:92148149
A:Accession: A46483
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-355 <LEF>
A:Cross-references: GB:M82873; NID:g165440; PID:g165441
A:Experimental source: neutrophils
A:Note: sequence extracted from NCBI backbone (NCBI:81526, NCBI:81530)
C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 2.8%; Score 10; DB 2; Length 355;
Best Local Similarity 100.0%; Pred. No. 0.035;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 133 DRYLAIVHAT 142
|||||
Db 139 DRYLAIVHAT 148

RESULT 9

Interleukin-8 receptor type B - mouse
C:Species: Mus musculus (house mouse)
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 17-Mar-1999
C:Accession: A48921; A53677; I49348; I5421; H48909; I53774
R:Cerretti, D.P.; Nelson, N.; Kozlosky, C.J.; Morrissey, P.J.; Copeland, N.G.; Gilber
Genomics 18, 410-413, 1993
A:Title: The murine homologue of the human interleukin-8 receptor type B maps near th
A:Reference number: A48921; MUID:94117014
A:Accession: A48921
A:Molecule type: DNA
A:Residues: 1-359 <CER>
A:Cross-references: GB:L23637; NID:g435093; PID:g435094
R:Suzuki, H.; Prado, G.N.; Wilkinson, N.; Navarro, J.
J. Biol. Chem. 269, 18263-18266, 1994
A:Title: The N terminus of interleukin-8 (IL-8) receptor confers high affinity bindin
A:Reference number: A53677; MUID:94308043
A:Accession: A53677
A:Molecule type: DNA
A:Residues: 1-359 <SUZ>
A:Cross-references: GB:L26549
A:Note: sequence extracted from NCBI backbone (NCBI:149812)
R:Lee, J.; Caclalano, G.; Camerato, T.; Toy, K.; Moore, M.W.; Wood, W.I.
J. Immunol. 159, 2158-2164, 1995
A:Title: Chemokine binding and activities mediated by the mouse IL-8 receptor.
A:Reference number: I49348; MUID:95363183
A:Accession: I49348
A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA
 A:Residues: 1-359 <RES>
 A:Cross-references: EMBL:U01207; NID:g950174; PID:g950175
 R:Bozic, C.R.; Gerard, N.P.; von Uexkull-Guldenband, C.; Kolakowski, L.F.
 J. Biol. Chem. 269, 29355-29358, 1994
 A:Title: The murine interleukin 8 type B receptor homologue and its ligands. Expression
 A:Reference number: 155421; MUID:95050766
 A:Accession: 155421
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-359 <RES>
 A:Cross-references: GB:J13339; NID:g293665; PID:g293666
 R:Wilkie, T.M.; Chen, Y.; Glibbert, D.J.; Moore, K.J.; Yu, L.; Simon, M.I.; Copeland, N.G.
 Genomics 18, 175-184, 1993
 A:Title: Identification, chromosomal location, and genome organization of mammalian G-protein coupled receptor 10
 A:Reference number: A48909; MUID:94116980
 A:Accession: H48909
 A:Status: preliminary; nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 145-258 <MIL>
 A:Cross-references: GB:J20337; NID:g438800; PID:g438801
 R:Harada, A.; Kuno, K.; Nomura, H.; Mukaida, N.; Murakami, S.; Matsushima, K.
 Gene 142, 297-300, 1994
 A:Title: Cloning of a cDNA encoding a mouse homolog of the interleukin-8 receptor.
 A:Reference number: 153774; MUID:94252584
 A:Accession: 153774
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-359 <RES>
 A:Cross-references: GB:DJ7630; NID:g493671; PID:g493672
 A:Gene: 118-b
 A:Genetics:
 A:Introns: #status absent
 C:Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein
 F:49-74/Domain: transmembrane #status predicted <TM1>
 F:84-106/Domain: transmembrane #status predicted <TM2>
 F:120-141/Domain: transmembrane #status predicted <TM3>
 F:163-182/Domain: transmembrane #status predicted <TM4>
 F:213-234/Domain: transmembrane #status predicted <TM5>
 F:251-271/Domain: transmembrane #status predicted <TM6>
 F:308-328/Domain: transmembrane #status predicted <TM7>

Query Match 2.8%; Score 10; DB 2; Length 359;
 Best Local Similarity 100.0%; Pred. No. 0.036;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 133 DRYLATVHAT 142
 |||||
 DB 142 DRYLATVHAT 151

RESULT 10
 K3 opiate receptor - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 02-Jul-1996
 C:Accession: J49022
 R:Pan, Y.X.; Cheng, J.; Xu, J.; Rossi, G.; Jacobson, E.; Ryan-Moro, J.; Brooks, A.I.; De Mol. Pharmacol. 47, 1180-1188, 1995
 A:Title: Cloning and functional characterization through antisense mapping of a kappa 3-opioid receptor
 A:Reference number: J49022; MUID:95327076
 A:Accession: J49022
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-367 <RES>
 A:Cross-references: EMBL:U09421; NID:g551484; PID:g551485

Query Match 2.8%; Score 10; DB 2; Length 367;
 Best Local Similarity 100.0%; Pred. No. 0.036;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 296 CLNPITLYAFL 305
 |||||
 DB 310 CLNPITLYAFL 319

RESULT 11
 JC2421
 oploid receptor homolog, MOR-C - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 21-Feb-1995 #sequence_revision 05-Apr-1995 #text_change 13-Mar-1997
 C:Accession: JC2421; J49122
 R:Nishi, M.; Takeshima, H.; Mori, M.; Nakagawara, K.; Takeuchi, T.
 Biochem. Biophys. Res. Commun. 205, 1353-1357, 1994
 A:Title: Structure and chromosomal mapping of genes for the mouse kappa-opioid receptor
 A:Reference number: JC2434; MUID:95100967
 A:Accession: JC2421
 A:Molecule type: mRNA
 A:Residues: 1-367 <NIS>
 A:Cross-references: DDBJ:D31663
 R:Halford, W.P.; Gebhardt, B.M.; Carr, D.J.
 J. Neuroimmunol. 59, 91-101, 1995
 A:Title: Functional role and sequence analysis of a lymphocyte orphan oploid receptor
 A:Reference number: J49122; MUID:95318231
 A:Accession: J49122
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-357 <RES>
 A:Cross-references: EMBL:U14165; NID:g540092; PID:g540093
 C:Genetics:
 A:Gene: MOR-C; OOR
 A:Map position: 2H2-4
 A:Introns: 75/2; 194/1
 C:Keywords: receptor

Query Match 2.8%; Score 10; DB 2; Length 367;
 Best Local Similarity 100.0%; Pred. No. 0.036;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 296 CLNPITLYAFL 305
 |||||
 DB 310 CLNPITLYAFL 319

RESULT 12
 G protein-coupled receptor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 29-Jan-1999
 C:Accession: I56520; S46238; S43655; S45701; I52654
 R:Lachowicz, J.E.; Shen, Y.; Monsma, F.J.
 J. Neurochem. 64, 34-40, 1995
 A:Title: Molecular cloning of a novel G protein-coupled receptor related to the opiat
 A:Reference number: I56520; MUID:95096849
 A:Accession: I56520
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-367 <RES>
 A:Cross-references: EMBL:U07871; NID:g606802; PID:g606803
 R:Chen, Y.; Fan, Y.; Liu, J.; Mestek, A.; Tian, M.; Kozak, C.A.; Yu, L.
 FEBS Lett. 347, 279-283, 1994
 A:Title: Molecular cloning, tissue distribution and chromosomal localization of a nov
 A:Reference number: S46238; MUID:94307400
 A:Accession: S46238
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-367 <CHE>
 A:Cross-references: GB:J28144; NID:g496219; PID:g496220
 R:Fukuda, K.; Kato, S.; Mori, K.; Nishi, M.; Takeshima, H.; Iwabe, N.; Miyata, T.; Ho
 FEBS Lett. 343, 42-46, 1994
 A:Title: cDNA cloning and regional distribution of a novel member of the oploid recep
 A:Reference number: S43655; MUID:94215703
 A:Accession: S43655

A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-367 <EUK>
 A:Cross-references: GB:D16438; NID:g456168; PID:Q1004424; PID:g533355
 R:Bunzow, J.R.; Saez, C.; Mottlind, M.; Bouvier, C.; Williams, J.T.; Low, M.; Grandy, D.K.
 FEBS Lett. 347, 284-288, 1994
 A:Title: Molecular cloning and tissue distribution of a putative member of the rat opioi
 A:Reference number: S46235; MUID:94307401
 A:Accession: S46235
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-347, 'A', 349-367 <BUN>
 A:Cross-references: EMBL:U01913; NID:g487964; PID:g487965
 A:Note: the authors translated the codon GCG for residue 348 as Ser
 R:Wang, J.B.; Johnson, P.S.; Imel, Y.; Persico, A.M.; Ozenberger, B.A.; Eppler, C.M.; UH
 FEBS Lett. 348, 75-79, 1994
 A:Title: cDNA cloning of an orphan opiate receptor gene family member and its splice var
 A:Reference number: S45701; MUID:94298959
 A:Accession: S45701
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-331, 'LL', 334-367 <WAN>
 R:Wick, M.J.; Minnerath, S.R.; Lin, X.; Elide, R.; Law, P.Y.; Loh, H.H.
 Brain Res. Mol. Brain Res. 27, 37-44, 1994
 A:Title: Isolation of a novel cDNA encoding a putative membrane receptor with high homol
 A:Reference number: I52654; MUID:95182817
 A:Accession: I52654
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-367 <RE2>
 A:Cross-references: GB:I29419; NID:g510718; PID:g510719
 C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 2.8%; Score 10; DB 2; Length 367;
 Best Local Similarity 100.0%; Pred. No. 0.036;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 296 CUNPILYAF 305
 |||||
 DB 310 CUNPILYAF 319

RESULT 13
 S42096
 Interleukin-8 receptor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Feb-1997
 C:Accession: S42096
 R:GOBL, A.E.; Wang, S.; Zhou, Y.; Oeberg, K.
 submitted to the EMBL Data Library, February 1994
 A:Description: Molecular cloning of the rat IL8 receptor.
 A:Reference number: S42096
 A:Accession: S42096
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-356 <GOB>
 A:Cross-references: EMBL:X77797
 C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 2.8%; Score 10; DB 2; Length 356;
 Best Local Similarity 100.0%; Pred. No. 0.035;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 133 DRYLAIVHAT 142
 |||||
 DB 142 DRYLAIVHAT 151

Search completed: August 8, 1999, 03:21:13
 Job time: 1307 sec

Thu Aug 12 09:17:08 1999

us-09-104-063-4.rpt

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 8, 1999, 05:59:13 ; Search time 10.2 Seconds
(without alignments)
975,533 Million cell updates/sec

Title: US-09-104-063-4

Score: 352

Sequence: 1 MEGSITYSDMYTEEMSGD.....KRGHSSVSTESSESSFHSS 352

Scoring table: OLIGO

Searched: 77977 seqs, 28258293 residues

Database : SwissProt_37.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	352	100.0	352	1	CCR4_HUMAN
2	213	60.5	352	1	CCR4_PAPAN
3	168	47.7	352	1	CCR4_CERIO
4	138	39.2	352	1	CCR4_MACMU
5	114	32.4	352	1	CCR4_MACRA
6	74	21.0	353	1	CCR4_BOVIN
7	74	21.0	353	1	CCR4_FELCA
8	60	17.0	359	1	CCR4_MOUSE
9	60	17.0	349	1	CCR4_SHEEP
10	55	15.6	192	1	CCR4_RAT
11	10	2.8	350	1	CCR4_SHEEP
12	10	2.8	350	1	CCR4_SHEEP
13	10	2.8	350	1	CCR4_SHEEP
14	10	2.8	350	1	CCR4_SHEEP
15	10	2.8	350	1	CCR4_SHEEP
16	10	2.8	350	1	CCR4_SHEEP
17	10	2.8	350	1	CCR4_SHEEP
18	10	2.8	350	1	CCR4_SHEEP
19	10	2.8	350	1	CCR4_SHEEP
20	10	2.8	350	1	CCR4_SHEEP
21	10	2.8	350	1	CCR4_SHEEP
22	10	2.8	350	1	CCR4_SHEEP
23	10	2.8	350	1	CCR4_SHEEP
24	10	2.8	350	1	CCR4_SHEEP
25	10	2.8	350	1	CCR4_SHEEP
26	10	2.8	350	1	CCR4_SHEEP
27	10	2.8	350	1	CCR4_SHEEP
28	10	2.8	350	1	CCR4_SHEEP

ALIGNMENTS

RESULT 1
ID CCR4_HUMAN STANDARD: PRT: 352 AA.
AC P30991: P56438: ..
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE C-X-C CHEMOKINE RECEPTOR TYPE 4 (CXCR-4) (CXCR-4) (SDF-1 RECEPTOR)
DE (STROMAL CELL-DERIVED FACTOR 1 RECEPTOR) (FUSIN) (LEUKOCYTE-DERIVED
SEVEN TRANSMEMBRANE DOMAIN RECEPTOR) (LESTR) (LCRI) (FB22) (NPYRL)

DE (HM89).
OS CXCR4.
OS HOMO SAPIENS (HUMAN), AND PAN TROGLODITES (CHIMPANZEE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES-HUMAN: TISSUE-LUNG;
RX MEDLINE: 93319629.
RA HEROG H., HORT Y.J., SHINE J., SELBIE L.A.;
RT "Molecular cloning, characterization, and localization of the human
RT homolog to the reported bovine NPY Y3 receptor: lack of NPY binding
RT and activation."
RL DNA CELL BIOL. 12:465-471(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES-HUMAN: TISSUE-BRAIN;
RX MEDLINE: 94052833.
RA JAZIN E.E., YOO H., BLOMQUIST G., YEE F., WENG G., WALKER M.W.,
RA SALON J., LARHAMMAR D., WAHLESTEDT C.R.;
RT "A proposed bovine neuropeptide Y (NPY) receptor cDNA clone, or its
RT human homologue, confers neither NPY binding sites nor NPY
RT responsiveness on transfected cells."
RL REGUL. PEPT. 47:247-258(1993).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES-HUMAN: TISSUE-SPLEEN;
RX MEDLINE: 93315164.
RA FEDERSPIEL B., DELANEY A.D., CLARK-LEWIS I., JIRIK F., DUNCAN A.M.,
RA SCHAPPERT K.T., MELHADO I.;
RT "Molecular cloning of the cDNA and chromosomal localization of the
RT gene for a putative seven-transmembrane segment (7-TMS) receptor
RT isolated from human spleen."
RL GENOMICS 16:707-712(1995).
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES-HUMAN: TISSUE-LEUKOCYTE;
RX MEDLINE: 94103215.
RA LOETSCHER M., GELSER T., O'REILLY T., ZWAHLER R., BAGGIOLINI M.,
RA MOSER B.;
RT "Cloning of a human seven-transmembrane domain receptor, LESTR, that
RT is highly expressed in leukocytes."
RL J. BIOL. CHEM. 269:232-237(1994).
RN [5]
RP SEQUENCE FROM N.A.
RC SPECIES-HUMAN: TISSUE-MONOCYTES;
RX MEDLINE: 94092629.
RA NOMURA H., NIELSEN B.W., MATSUSHIMA K.;
RT "Molecular cloning of cDNA encoding a LD78 receptor and putative
RT leukocyte chemotactic peptide receptors."
RL INT. IMMUNOL. 5:1239-1249(1993).
RN [6]
RP SEQUENCE FROM N.A.
RC SPECIES-HUMAN;
RX WEGNER S.A., EHRENBERG P.K., CHANG G., DAYHOFF D.E., MICHAEL N.L.;
RL SUBMITTED (MAY-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [7]
RP SEQUENCE FROM N.A.
RC SPECIES-P. TROGLODITES;
RX MEDLINE: 98090115.
RA PRETET J.-L., ZERBIB A., GIRARD M., GUILLET J.-G., BUTOR C.;
RT "Chimpanzee CXCR4 and CCR5 act as coreceptors for HIV type 1."
RL AIDS RES. HUM. RETROVIRUSES 13:1583-1587(1997).
RN [8]
RP FUNCTION.
RX MEDLINE: 96351077.
RA BLEUL C.C., FARZAN M., CHOE H., PAROLIN C., CLARK-LEWIS I.,
RA SODROSKI J., SPRINGER T.A.;
RT "The lymphocyte chemotactant SDF-1 is a ligand for LESTR/fusin and
RT blocks HIV-1 entry."
RL NATURE 382:829-833(1996).
RN [9]
RP FUNCTION.

```

FT CARBOHYD 11 11 POTENTIAL.
FT DISUFED 109 186 BY SIMILARITY.
SQ SEQUENCE 352 AA: 39745 MW: BDE02133 CRC32:

Query Match 100.0%; Score 352; DB 1; Length 352:
Best Local Similarity 100.0%; Pred. No. 0;
Matches 352; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEGISITSDNYTEEMSSGDYDSMKKECFEENANNEKIFLPITYSIPLTGIVGGLVI 60
Db 1 MEGISITSDNYTEEMSSGDYDSMKKECFEENANNEKIFLPITYSIPLTGIVGGLVI 60
QY 61 LVMGYOKKLRSMETKRYLHLTVADLFLVITLPEWADAVANWYFGNFLCAVHYIYTVNL 120
Db 61 LVMGYOKKLRSMETKRYLHLTVADLFLVITLPEWADAVANWYFGNFLCAVHYIYTVNL 120
QY 121 YSSVLLIAFLSDRYLAIVAHATNSQRPRLKLAENVYVGVWIPALLITIDPFANSEA 180
Db 121 YSSVLLIAFLSDRYLAIVAHATNSQRPRLKLAENVYVGVWIPALLITIDPFANSEA 180
QY 181 DDRYICDAFYENDLVWVYFQFHLMVGHILPGIYLISCYCIILSKLSHGOKRKALKT 240
Db 181 DDRYICDAFYENDLVWVYFQFHLMVGHILPGIYLISCYCIILSKLSHGOKRKALKT 240
QY 241 TVILLIAEFACMLPYYIGISIDSFILLEIKQGEFENVYHAKMISTIDALAFHCCLNPI 300
Db 241 TVILLIAEFACMLPYYIGISIDSFILLEIKQGEFENVYHAKMISTIDALAFHCCLNPI 300
QY 301 LYAFIAGAFKTSAGHALTVSRGSSKILTSKRGKGHSVSTESSESSFHSS 352
Db 301 LYAFIAGAFKTSAGHALTVSRGSSKILTSKRGKGHSVSTESSESSFHSS 352

RESULT 2
ID CCR4.PAPAN STANDARD: PRT: 352 AA.
AC P56491.
DT 15-JUL-1998 (REL. 36, CREATED)
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DI 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE C-X-C CHEMOKINE RECEPTOR TYPE 4 (CXCR-4) (SDF-1 RECEPTOR)
DE (STROMAL CELL-DERIVED FACTOR 1 RECEPTOR) (FUSIN).
GN CXCR4.
OS PAP10 ANUBIS (OLIVE BABOON).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; CERCOPITHECIDAE; CERCOPITHECINAE; PAP10.
RN [1]
RP SEQUENCE FROM N.A.
RA BENION P.A., TIMANUS D.K., SHEARER M.H., LEE D.R., KENNEDY R.C.;
RL SUBMITTED (NOV-1997) TO EMBL/GENBANK/DDJ DATA BANKS.
CC -1- FUNCTION: RECEPTOR FOR THE C-X-C CHEMOKINE SDF-1. TRANSUCES A
CC SIGNAL BY INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL.
CC -1- SUBCELLULAR LOCATION: INTERAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF031089; G2625094; -
DR GCRDB: GCR_2512; -
DR PROSITE: PS00237; G_PROTEIN_RECEPTOR; 1.
KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRAN; GLYCOPROTEIN.
FT DOMAIN 1 39
FT TRANSMEM 40 63
FT DOMAIN 64 79
FT TRANSMEM 80 99
FT DOMAIN 100 110
EXTRACELLULAR (POTENTIAL).

```



```

FT TRANSMEM 111 132 3 (POTENTIAL).
FT DOMAIN 133 134 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 155 175 4 (POTENTIAL).
FT DOMAIN 176 200 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 201 220 5 (POTENTIAL).
FT DOMAIN 221 240 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 241 261 6 (POTENTIAL).
FT DOMAIN 262 285 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 286 305 7 (POTENTIAL).
FT DOMAIN 306 352 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 11 11 POTENTIAL.
FT DISULFID 109 186 BY SIMILARITY.
SQ SEQUENCE 352 AA: 39751 MM: 273DB8EE CRC32:

Query Match
Best Local Similarity 60.5%; Score 213; DB 1: Length 352;
Matches 313; Conservative 0; Mismatches 1; Indels 0; Gaps 0:

QY 39 ILEPTIYSITFLGIYNGVILVMGYQKRLRSMTDKYRLHLSVADLLFVITLPEFAVDA 98
DB 39 ILEPTIYSITFLGIYNGVILVMGYQKRLRSMTDKYRLHLSVADLLFVITLPEFAVDA 98
QY 99 VANWYFGNFKRAVHYIYVNLKXSVLILAFISDRYLAVHATNSORPKLLAEKVVYV 158
DB 99 VANWYFGNFKRAVHYIYVNLKXSVLILAFISDRYLAVHATNSORPKLLAEKVVYV 158
QY 159 GWAIPALLLTIPDFIFANVSEADRYICDRFPYNDLWVWVFOFOHIMVGLILPGIVILSC 218
DB 159 GWAIPALLLTIPDFIFANVSEADRYICDRFPYNDLWVWVFOFOHIMVGLILPGIVILSC 218
QY 219 YCIIISKLSKSGHOKRAKTKTIVILILAFACMLPYIGISIDSFTLLIINOGCEFFEN 278
DB 219 YCIIISKLSKSGHOKRAKTKTIVILILAFACMLPYIGISIDSFTLLIINOGCEFFEN 278
QY 279 TVHKWISITELALFFHCCLNPILYAFILGAKFKTSAGQHALTSVSRGSSSLKILSKGRGHS 338
DB 279 TVHKWISITELALFFHCCLNPILYAFILGAKFKTSAGQHALTSVSRGSSSLKILSKGRGHS 338
QY 339 SVSTESSESSSFHSS 352
DB 339 SVSTESSESSSFHSS 352

RESULT 3
CCR4_CERTO
ID CCR4_CERTO STANDARD: PRT: 352 AA.
AC 062747:
DT 15-DEC-1998 (REL. 37, CREATED)
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE C-X-C CHEMOKINE RECEPTOR TYPE 4 (CXCR-4) (SDF-1 RECEPTOR)
DE (STROMAL CELL-DERIVED FACTOR 1 RECEPTOR) (FUSIN) (LESTR).
GN CXCR4.
OS CEROCCEBUS TOROUATUS AIYS (RED-CROWNED MANGABEY) (SOOTY MANGABEY).
OC EUKARYOTA: METAZOA: CHORDATA: VERTEBRATA: MAMMALIA: EUTHERIA:
OC PRIMATES: CATARRHINI: CERCOPIITHECIDAE: CERCOPIITHECINAE: CERCOCEBUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 98321155.
RA CHEN Z., GETTIE A., HO D.D., MARX P.A.:
"Primary stvsm isolates use the CCR5 coreceptor from sooty mangabeys
naturally infected in west Africa: a comparison of coreceptor usage
of primary stvsm. HIV-2, and stvmac."
VIROLOG 246:113-124(1998).
RL
CC -1- FUNCTION: RECEPTOR FOR THE C-X-C CHEMOKINE SDF-1. TRANSDUCES A
CC SIGNAL BY INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF051906; G3135304; -
DR PROSITE: PS00237; G_PROTEIN_RECEPTOR; 1.
KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN.
FT DOMAIN 1 39
FT TRANSMEM 40 63 1 (POTENTIAL).
FT DOMAIN 64 79 2 (POTENTIAL).
FT TRANSMEM 80 99 3 (POTENTIAL).
FT DOMAIN 100 110 4 (POTENTIAL).
FT TRANSMEM 111 132 5 (POTENTIAL).
FT DOMAIN 133 154 6 (POTENTIAL).
FT TRANSMEM 155 175 7 (POTENTIAL).
FT DOMAIN 176 200 8 (POTENTIAL).
FT TRANSMEM 201 220 9 (POTENTIAL).
FT DOMAIN 221 240 10 (POTENTIAL).
FT TRANSMEM 241 261 11 (POTENTIAL).
FT DOMAIN 262 285 12 (POTENTIAL).
FT TRANSMEM 286 305 13 (POTENTIAL).
FT DOMAIN 306 352 14 (POTENTIAL).
FT CARBOHYD 11 11 POTENTIAL.
FT DISULFID 109 186 BY SIMILARITY.
SQ SEQUENCE 352 AA: 39648 MM: 8B464DAD CRC32:

Query Match
Best Local Similarity 47.7%; Score 168; DB 1: Length 352;
Matches 168; Conservative 0; Mismatches 0; Indels 0; Gaps 0:

QY 185 ICDRFYNDLWVWVFOFOHIMVGLILPGIVILSCYCIISKLSKSGHOKRAKTKTIVIL 244
DB 185 ICDRFYNDLWVWVFOFOHIMVGLILPGIVILSCYCIISKLSKSGHOKRAKTKTIVIL 244
QY 245 ILAFACMLPYIGISIDSFTLLIETIKOGCEFFENTYHKWISITELALFFHCCLNPILYAF 304
DB 245 ILAFACMLPYIGISIDSFTLLIETIKOGCEFFENTYHKWISITELALFFHCCLNPILYAF 304
QY 305 LGAKFKTSAGQHALTSVSRGSSSLKILSKGRGHSVSTESSESSSFHSS 352
DB 305 LGAKFKTSAGQHALTSVSRGSSSLKILSKGRGHSVSTESSESSSFHSS 352

RESULT 4
CCR4_MACMU
ID CCR4_MACMU STANDARD: PRT: 352 AA.
AC P79394; 002745; 046428;
DT 15-JUL-1998 (REL. 36, CREATED)
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE C-X-C CHEMOKINE RECEPTOR TYPE 4 (CXCR-4) (SDF-1 RECEPTOR)
DE (STROMAL CELL-DERIVED FACTOR 1 RECEPTOR) (FUSIN) (LESTR).
GN CXCR4.
OS MACACA MULATTA (RHESUS MACAQUE).
OC EUKARYOTA: METAZOA: CHORDATA: VERTEBRATA: MAMMALIA: EUTHERIA:
OC PRIMATES: CATARRHINI: CERCOPIITHECIDAE: CERCOPIITHECINAE: MACACA.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-INDIAN MACAQUE;
RX MEDLINE: 97213934.
RA CHEN Z., ZHOU P., HO D.D., LANDAU N.R., MARX P.A.:
"Genetically divergent strains of simian immunodeficiency virus use
CCR5 as a coreceptor for entry."
J. VIROL. 71:2705-2714(1997).
RL [2]
CC SEQUENCE FROM N.A.
CC MEDLINE: 97268687.
CC EDINGER A.L., AMEDEE A., MILLER K., DORANZ B.J., ENDRES M.,
CC SHARON M., SAMSON M., LU Z.-H., CLEMENTS J.E., MORPHEE-CORB M.,
CC PEIPER S.C., PARMENTIER M., BRODER C.C., DONS R.W.:
CC "Differential utilization of CCR5 by macrophage and T cell tropic

```

```

RT simian immunodeficiency virus strains".
RL PROC. NATL. ACADE. SCI. U.S.A. 94:4005-4010(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=CHINESE:
RA MEDLINE: 98252393.
RA PRETER J.-L., GUILLET J.-G., BUTOR C.:
RT "New widespread CXCR4 allele in rhesus macaques does not predict
RT subspecies or clinical evolution."
RL AIDS RES. HUM. RETROVIRUSES 14:639-641(1998).
CC -1- FUNCTION: RECEPTOR FOR THE C-X-C CHEMOKINE SDF-1. TRANSDUCES A
CC SIGNAL BY INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U73740; G1899056; -
CC EMBL: U93311; G1934671; -
CC EMBL: AF001928; G2911294; -
CC GCRDB: GCR.1297; -
CC GCRDB: GCR.1640; -
CC PROSITE: PS00237; G-PROTEIN_RECEPTOR; 1.
CC PFAM: PF00001; 7tm_1; 1.
DR G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN.
FT DOMAIN 1 39
FT TRANSMEM 40 63
FT DOMAIN 64 79
FT TRANSMEM 80 99
FT DOMAIN 100 110
FT TRANSMEM 111 132
FT DOMAIN 133 154
FT TRANSMEM 155 175
FT DOMAIN 176 200
FT TRANSMEM 201 220
FT DOMAIN 221 240
FT TRANSMEM 241 261
FT DOMAIN 262 285
FT TRANSMEM 286 305
FT DOMAIN 306 352
FT CARBOHYD 11
FT DISULFID 109 186
FT CONFLICT 67 67
FT CONFLICT 214 214
FT CONFLICT 348 348
SQ SEQUENCE 352 AA: 39739 MW: 3C33A4BB CRC32:

Query Match 39.2%; Score 138; DB 1; Length 352;
Best Local Similarity 100.0%; Pred. No. 2.4e-131;
Matches 138; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

ID CCR4_MACEFA STANDARD: PRT: 352 AA.
AC 028474;
DT 15-JUL-1998 (REL. 36, CREATED)
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE C-X-C CHEMOKINE RECEPTOR TYPE 4 (CXCR-4) (CXCR-4) (SDF-1 RECEPTOR)
DE (STROMAL CELL-DERIVED FACTOR 1 RECEPTOR) (FUSIN) (LESTR).
GN CXCR4.
OS MACACA FASCICULARIS (CRAB EATING MACAQUE) (CYNOMOLGUS MONKEY).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; CERCOPITHECIDAE; CERCOPITHECINAE; MACACA.
RN [1]
RP SEQUENCE FROM N.A.
RA TATSUMI M., TAKAHASHI H.;
RT SUBMITTED (JUL-1996) TO EMBL/GENBANK/DDSI DATA BANKS.
RL
CC -1- FUNCTION: RECEPTOR FOR THE C-X-C CHEMOKINE SDF-1. TRANSDUCES A
CC SIGNAL BY INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D86579; G1468949; -
CC GCRDB: GCR.1143; -
CC PROSITE: PS00237; G-PROTEIN_RECEPTOR; 1.
CC PFAM: PF00001; 7tm_1; 1.
DR G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN.
FT DOMAIN 1 39
FT TRANSMEM 40 63
FT DOMAIN 64 79
FT TRANSMEM 80 99
FT DOMAIN 100 110
FT TRANSMEM 111 132
FT DOMAIN 133 154
FT TRANSMEM 155 175
FT DOMAIN 176 200
FT TRANSMEM 201 220
FT DOMAIN 221 240
FT TRANSMEM 241 261
FT TRANSMEM 262 285
FT TRANSMEM 286 305
FT DOMAIN 306 352
FT CARBOHYD 11
FT DISULFID 109 186
SQ SEQUENCE 352 AA: 39753 MW: 7EDA93BA CRC32:

Query Match 32.4%; Score 114; DB 1; Length 352;
Best Local Similarity 100.0%; Pred. No. 3.7e-107;
Matches 114; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

RESULT 5
CCR4_MACEFA

RESULT 6
CCR4_BOVIN
ID CCR4_BOVIN STANDARD: PRT: 353 AA.
AC P25930;
DT 01-MAY-1992 (REL. 22, CREATED)
DT 01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)

DE C-X-C CHEMOKINE RECEPTOR TYPE 4 (CXCR-4) (SDF-1 RECEPTOR)
DE (STROMAL CELL-DERIVED FACTOR 1 RECEPTOR) (FUSIN) (LEUKOCYTE-DERIVED
DE SEVEN TRANSMEMBRANE DOMAIN RECEPTOR) (LESTR) (LCR1).
GN CXCR4.
OS BOS TAURUS (BOVINE).
OC EURARVOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC ARTIODACTYLA; RUMINANTIA; PECORA; BOVIDEA; BOVIAE; BOVINE; BOS.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-LOCUS COERULEUS.
RA MEDLINE: 92100053.
RX SWELAND J., XIN W., SWEETNAM P., SAJOH K., NESTLER E.J., DUMAN R.S.;
RL "Sequence and expression of a neuropeptide y receptor cDNA.";
RL MOL. PHARMACOL. 40:869-875(1991).
RN [2]
RN SHOWS THAT IT IS NOT A NPY3-R.
RX MEDLINE: 94052833.
RA JAZIN E.E., YOO H., BLOMOVIST G., YEE F., WENG G., WALKER M.W.,
RA SALON J., LARHAMMAR D., WAHLESTEDT C.R.;
RL "A proposed bovine neuropeptide y (NPY) receptor cDNA clone, or its
RL human homologue, confers neither NPY binding sites nor NPY
RT responsiveness on transfected cells.";
RL REGUL. PEPT. 47:247-258(1993).
CC -1- FUNCTION: RECEPTOR FOR THE C-X-C CHEMOKINE SDF-1. TRANSDUCES A
CC SIGNAL BY INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: BRAIN, HEART, KIDNEY, LUNG AND LIVER.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -1- CAUTION: WAS ORIGINALLY (REF.1) THOUGHT TO BE A RECEPTOR FOR
CC NEUROPEPTIDE Y, TYPE 3 (NPY3-R).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M86739; -; NOT_ANNOTATED_CDS.
DR PIR: S28787; S28787.
DR GCRDB: GCR_0180; -;
DR PROSITE: PS00237; G-PROTEIN_RECEPTOR; 1.
DR PFAM: PF00001; 7tm_1; 1.
KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN.
FT DOMAIN 1 40
FT TRANSMEM 41 64
FT DOMAIN 65 80
FT TRANSMEM 81 100
FT DOMAIN 101 111
FT TRANSMEM 112 133
FT DOMAIN 134 155
FT TRANSMEM 156 176
FT DOMAIN 177 201
FT TRANSMEM 202 221
FT DOMAIN 222 241
FT TRANSMEM 242 262
FT DOMAIN 263 286
FT TRANSMEM 287 306
FT DOMAIN 307 333
FT CARBOHYD 11 11
FT DISULFID 110 187
FT SEQUENCE 353 AA; 39938 MW; 03D8F100 CRC32;
Query Match 21.0%; Score 74; DB 1; Length 353;
Best Local Similarity 100.0%; Pred. No. 7; 7e-67;
Matches 74; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 279 TVHKWISITLALAFHCCPLILYAFLGAKFKTSQAHLTVSVSGSSLKILSKGRGHS 338
DB 280 TVHKWISITLALAFHCCPLILYAFLGAKFKTSQAHLTVSVSGSSLKILSKGRGHS 339

QY 339 SVSTESSSSSFHSS 352
DB 340 SVSTESSSSSFHSS 353
RESULT 7
ID CCR4_FELCA STANDARD; PRT: 353 AA.
AC P56498; F79172; 002700;
DT 15-JUL-1998 (REL. 36, CREATED)
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE C-X-C CHEMOKINE RECEPTOR TYPE 4 (CXCR-4) (SDF-1 RECEPTOR)
DE (STROMAL CELL-DERIVED FACTOR 1 RECEPTOR) (FUSIN) (LESTR).
GN CXCR4.
OS FELIS SILVESTRIS CATUS (CAT).
OC EURARVOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC CARNIVORA; FISSIPEDIA; FELIDAE; FELIS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 97404646.
RA WILLETT B.J., PICARD L., HOSIE M.J., TURNER J.D., ADEMA K.,
RA CLAPHAM P.R.;
RL "Shared usage of the chemokine receptor CXCR4 by the feline and human
RL immunodeficiency viruses.";
RL J. VIROL. 71:6407-6415(1997).
RN [2]
RN SEQUENCE FROM N.A.
RA WILLETT B.J.;
RL SUBMITTED (DEC-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RA LERNER D.L., ELDER J.H.;
RL SUBMITTED (MAR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- FUNCTION: RECEPTOR FOR THE C-X-C CHEMOKINE SDF-1. TRANSDUCES A
CC SIGNAL BY INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U63558; G1842247; -;
DR EMBL: U92795; G1935045; -;
DR GCRDB: GCR_1113; -;
DR GCRDB: GCR_1114; -;
DR PROSITE: PS00237; G-PROTEIN_RECEPTOR; 1.
KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN.
FT DOMAIN 1 40
FT TRANSMEM 41 64
FT DOMAIN 65 80
FT TRANSMEM 81 100
FT DOMAIN 101 111
FT TRANSMEM 112 133
FT DOMAIN 134 155
FT TRANSMEM 156 176
FT DOMAIN 177 201
FT TRANSMEM 202 221
FT DOMAIN 222 241
FT TRANSMEM 242 262
FT DOMAIN 263 286
FT TRANSMEM 287 306
FT DOMAIN 307 353
FT CARBOHYD 11 11
FT DISULFID 110 187
FT CONFLICT 67 263
FT SEQUENCE 353 AA; 39935 MW; 5829DDIE CRC32;
O -> H (IN REF. 3).
D -> E (IN REF. 3).

Query Match 21.0%; Score 74; DB 1; Length 353;
 Best Local Similarity 100.0%; Pred. No. 7,7e-67;
 Matches 74; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 279 TVHKMISTEALAFHCCNPLTVAFLGAKFKTSQHALTSVRSGLSKLTKGKRGHS 338
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 280 TVHKMISTEALAFHCCNPLTVAFLGAKFKTSQHALTSVRSGLSKLTKGKRGHS 339
 Y 339 SVTSESSSESSFHSS 352
 ||||||||||||||||
 Db 340 SVTSESSSESSFHSS 353

RESULT 8

CCRA_MOUSE STANDARD; PRT: 359 AA.
 ID CCR4_MOUSE P70658; P70346: 009062; 009059; P70233;
 AC 01-NOV-1997 (REL. 35, CREATED)
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
 DE C-X-C CHEMOKINE RECEPTOR TYPE 4 (CXCR-4) (SDF-1 RECEPTOR)
 DE (STROMAL CELL-DERIVED FACTOR 1 RECEPTOR) (FUSIN) (LEUKOCYTE-DERIVED
 SEVEN TRANSMEMBRANE DOMAIN RECEPTOR) (LESTR).
 GN CXCR4 OR LESTR OR CMKAR4 OR SDF1R.
 OS MUS MUSCULUS (MOUSE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
 RN RODENTIA; SCIROGNATHI; MURIDAE; MURINAE; MOS.
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6 X CBA; TISSUE=THYMUS;
 RA MOEPFS B., FRODL R., KESSLER H., GIERSCHEK P.;
 RL SUBMITTED (NOV-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
 [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=129/SV;
 RC HESEN M., BERMAN M.A., GERARD C., DORF M.E.;
 RL SUBMITTED (SEP-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
 [3]
 RN SEQUENCE FROM N.A.
 RP STRAIN=C57BL/6J;
 RC MEDLINE: 97113334.
 RA HESEN M., BERMAN M.A., BENSON J.D., GERARD C., DORF M.E.;
 RT "Cloning of the mouse fusin gene, homologue to a human HIV-1
 co-factor";
 RT J. IMMUNOL. 157:5455-5460(1996).
 [4]
 RN SEQUENCE FROM N.A.
 RP TISSUE=BONE MARROW;
 RC MEDLINE: 97121456.
 RA NAGASAWA T., NAKAUMA T., TACHIBANA K., ITZASA H., BLEUL C.C.,
 RA YOSHIE O., MATSUSHIMA K., YOSHIDA N., SPRINGER T.A., KISHIMOTO T.;
 RT "Molecular cloning and characterization of a murine pre-B-cell
 growth-stimulating factor/stromal cell-derived factor 1 receptor, a
 murine homolog of the human immunodeficiency virus 1 entry coreceptor
 fusin";
 RT PROC. NATL. ACAD. SCI. U.S.A. 93:14726-14729(1996).
 [5]
 RN SEQUENCE FROM N.A.
 RP STRAIN=C57BL/6; TISSUE=THYMUS;
 RA SUZUKI G., NAKATA Y., UZAWA A., SHIRASAWA T., SAITO T., MITA K.;
 RL SUBMITTED (FEB-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 [6]
 RN SEQUENCE FROM N.A.
 RP STRAIN=129/SV; TISSUE=THYMUS;
 RA SCHUBEL A., BURGSTAHLER R., LIPP M.;
 RL SUBMITTED (SEP-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC -1- FUNCTION: RECEPTOR FOR THE C-X-C CHEMOKINE SDF-1. TRANSDUCES A
 SIGNAL BY INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: X99581; E281127; -;
 DR EMBL: X99582; E281126; -;
 DR EMBL: U59760; G1527135; -;
 DR EMBL: U65580; G1731651; -;
 DR EMBL: D87747; G1772446; -;
 DR EMBL: AB000803; G1816446; -;
 DR EMBL: 280111; E266602; -;
 DR EMBL: 280112; E181578; -;
 DR GCRDB; GCR_1138; -;
 DR GCRDB; GCR_1387; -;
 DR GCRDB; GCR_1646; -;
 DR GCRDB; GCR_1730; -;
 DR GCRDB; GCR_2592; -;
 DR MGD; MGI:109563; CMKAR4.
 DR PROSITE; PS00237; G_PROTEIN_RECEPTOR; 1.
 DR PFM; PF00001; 7tm_1; 1.
 KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN.
 FT DOMAIN 1 41
 FT TRANSMEM 42 65
 FT DOMAIN 66 81
 FT TRANSMEM 82 101
 FT DOMAIN 102 112
 FT TRANSMEM 113 134
 FT DOMAIN 135 156
 FT TRANSMEM 157 177
 FT DOMAIN 178 207
 FT TRANSMEM 208 227
 FT DOMAIN 228 247
 FT TRANSMEM 248 268
 FT DOMAIN 269 292
 FT TRANSMEM 293 312
 FT DOMAIN 313 359
 FT DISULFID 111 193
 FT CARBOHYD 13 13
 FT CONFLICT 6 7
 FT CONFLICT 216 216
 FT SEQUENCE 359 AA; 40426 MW; 1037B4D3 CRC32;

Query Match 17.0%; Score 60; DB 1; Length 359;
 Best Local Similarity 100.0%; Pred. No. 1e-52;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 197 VVFOFHIMWGLIPGIVILSCYCIITISKLSHSGOKRKALKTIVILLAFACWLPY 256
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 204 VVFOFHIMWGLIPGIVILSCYCIITISKLSHSGOKRKALKTIVILLAFACWLPY 263
 RESULT 9
 CCR4_RAT STANDARD; PRT: 349 AA.
 ID CCR4_RAT 008565;
 AC 008565;
 DT 01-NOV-1997 (REL. 35, CREATED)
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
 DE C-X-C CHEMOKINE RECEPTOR TYPE 4 (CXCR-4) (SDF-1 RECEPTOR)
 DE (STROMAL CELL-DERIVED FACTOR 1 RECEPTOR) (FUSIN) (LEUKOCYTE-DERIVED
 SEVEN TRANSMEMBRANE DOMAIN RECEPTOR) (LESTR).
 GN CXCR4 OR CMKAR4.
 OS RATTUS NOREUGIUS (RAT).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
 OC RODENTIA; SCIROGNATHI; MURIDAE; MURINAE; RATTUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=WISTAR; TISSUE=SPLEEN;

```

RA HARRISON J.K., SALAFRANCA M.N.:
RL SUBMITTED (MAR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- FUNCTION: RECEPTOR FOR THE C-X-C CHEMOKINE SDF-1. TRANSDUCES A
CC SIGNAL BY INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U90610; G1906613; -.
CC GCRDB: GCR_1401; -.
DR GCRDB: GCR_1401; -.
DR PROSITE: PS00237; G-PROTEIN_RECEPTOR; 1.
DR PFAM: PF00001; 7tm.1; 1.
KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN.
FT DOMAIN 1 36
FT TRANSMEM 37 60
FT DOMAIN 61 76
FT TRANSMEM 77 96
FT DOMAIN 97 107
FT TRANSMEM 108 129
FT DOMAIN 130 151
FT TRANSMEM 152 172
FT DOMAIN 173 197
FT TRANSMEM 198 217
FT DOMAIN 218 237
FT TRANSMEM 238 258
FT DOMAIN 259 282
FT TRANSMEM 283 302
FT DOMAIN 303 349
FT DISULFID 106 183
FT CARBOHYD 8
SQ SEQUENCE 349 AA; 39334 MW; CBA6532 CRC32;

Query Match 17.0%; Score 60; DB 1; Length 349;
Best Local Similarity 100.0%; Pred. No. 9.8e-53;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 197 VVEQFQHWGLILPGIVILSCYCIITISKLSHSGHOKRAKLTVTYLLIAFAACMPY 256
DB 194 VVFQFQHWGLILPGIVILSCYCIITISKLSHSGHOKRAKLTVTYLLIAFAACMPY 253

RESULT 10
CCR4_SHEEP STANDARD: PRT: 192 AA.
AC Q28553;
ID 15-JUL-1998 (REL. 36, CREATED)
DI 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE C-X-C CHEMOKINE RECEPTOR TYPE 4 (CXCR-4) (CXCR-4) (SDF-1 RECEPTOR)
DE (STROMAL CELL-DERIVED FACTOR 1 RECEPTOR) (FUSIN) (LEUCOCYTE-DERIVED
DE SEVEN TRANSMEMBRANE DOMAIN RECEPTOR) (LESTR) (FRAGMENT).
GN CXCR4 OR LESTR.
OS OVIS ARIES (SHEEP).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC ARTIODACTYLA; RUMINANTIA; PECORA; BOVIDAE; BOVIDAE; CAPRINAE; OVIS.
RN [1]
RP TISSUE-HYPOTHALAMUS;
RC DYER C.J., MATTERI R.L., KETSLER D.H.;
RI "Development of an ovine Y3 cDNA and expression of the Y3 receptor
RI mRNA in the ovine hypothalamus and pituitary.";
RL ABSTR. - SOC. NEUROSCI. 21:1890-1890(1995).
CC -1- FUNCTION: RECEPTOR FOR THE C-X-C CHEMOKINE SDF-1. TRANSDUCES A
CC SIGNAL BY INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.

```

```

CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -1- CAUTION: WAS ORIGINALLY THOUGHT TO BE A RECEPTOR FOR NEUROPEPTIDE
CC Y, TYPE 3 (NPY3-R).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U38942; G1061412; -.
CC GCRDB: GCR_1581; -.
DR GCRDB: GCR_1581; -.
DR PROSITE: PS00237; G-PROTEIN_RECEPTOR; 1.
DR PFAM: PF00001; 7tm.1; 1.
KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN.
FT NON_TER 1 1
FT DOMAIN <1 29
FT TRANSMEM 30 53
FT DOMAIN 54 69
FT TRANSMEM 70 89
FT DOMAIN 90 100
FT TRANSMEM 101 122
FT DOMAIN 123 144
FT TRANSMEM 145 165
FT DOMAIN 166 190
FT TRANSMEM 191 >192
FT DISULFID 99 176
FT NON_TER 192 192
SQ SEQUENCE 192 AA; 22178 MW; 6B3DEBD5 CRC32;

Query Match 15.6%; Score 55; DB 1; Length 192;
Best Local Similarity 100.0%; Pred. No. 6.3e-48;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 107 FLCKAVHYITVNLSSVLIATFISIDRYLAIVHANSQPRKLLKQVYGVW 161
DB 97 FLCKAVHYITVNLSSVLIATFISIDRYLAIVHANSQPRKLLKQVYGVW 151

RESULT 11
IL8A_GORGO STANDARD: PRT: 350 AA.
AC P55919; P55921;
DI 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE HIGH AFFINITY INTERLEUKIN-8 RECEPTOR A (IL-8R A) (IL-8 RECEPTOR TYPE
DE 1) (CXCR-1) (CDW128).
GN IL8RA OR CXCR1.
OS GORILLA GORILLA GORILLA (LOWLAND GORILLA), AND
OS PONGO PYGMAEUS (ORANGUTAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; GORILLA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 96175151.
RA ALVAREZ V., COTO E., SETIEN F., GONZALEZ S., GONZALEZ-ROCES S.,
RA LOPEZ-LARREA C.;
RI "Characterization of interleukin-8 receptors in non-human primates.";
RI IMMUNOGENETICS 43:261-267(1996).
CC -1- FUNCTION: RECEPTOR TO INTERLEUKIN-8, WHICH IS A POWERFUL
CC NEUTROPHILS CHEMOTACTIC FACTOR. BINDING OF IL-8 TO THE RECEPTOR
CC CAUSES ACTIVATION OF NEUTROPHILS. THIS RESPONSE IS MEDIATED VIA A
CC G-PROTEIN THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
CC MESSENGER SYSTEM. THIS RECEPTOR BINDS TO IL-8 WITH A HIGH AFFINITY
CC AND TO MGSA (GRO) WITH A LOW AFFINITY.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```


FT	TRANSEM	40	66	1	(POTENTIAL) .
FT	DOMAIN	67	75	2	CYTOPLASMIC (POTENTIAL) .
FT	TRANSEM	76	96	3	EXTRACELLULAR (POTENTIAL) .
FT	DOMAIN	97	111	4	(POTENTIAL) .
FT	TRANSEM	112	133	5	CYTOPLASMIC (POTENTIAL) .
FT	DOMAIN	134	154	6	(POTENTIAL) .
FT	TRANSEM	155	174	7	EXTRACELLULAR (POTENTIAL) .
FT	DOMAIN	175	199	8	(POTENTIAL) .
FT	TRANSEM	200	220	9	EXTRACELLULAR (POTENTIAL) .
FT	DOMAIN	220	240	10	(POTENTIAL) .

EMBL, LMB, EMBL (N-TERMINAL PEPTIDE RECEPTION).

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb.ch/announce/> or send an email to license@isb-sib.ch).

```
DR EMBL: M74240; G155439; -
DR EMBL: M58021; G155443; -
DR EMBL: M82873; G155441; -
DR PIR: A23669; A23669; -
DR PIR: A46483; A46483; -
DR PIR: J01231; J01231; -
DR GCRDB: GCR_0107; -
DR GCRDB: GCR_0108; -
DR GCRDB: GCR_0298; -
DR PROSITE: PS00237; G_PROTEIN_RECEPTOR; 1.
DR PFAM: PF00001; 7tm_1; 1.
KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN;
KW CHEMOTAXIS.
FT DOMAIN 1 40
FT TRANSSEM 41 67
FT DOMAIN 68 73
FT TRANSSEM 74 92
FT DOMAIN 93 114
FT TRANSSEM 115 138
FT TRANSSEM 139 159
FT TRANSSEM 160 184
FT TRANSSEM 185 204
FT TRANSSEM 205 232
FT TRANSSEM 233 247
FT TRANSSEM 248 270
FT TRANSSEM 271 290
FT TRANSSEM 291 313
FT TRANSSEM 314 355
FT CARBOHYD 7 21
FT CARBOHYD 21 192
FT DISULFID 90 111
FT CONFLICT 146 147
FT CONFLICT 204 204
FT CONFLICT 287 288
FT SEQUENCE 355 AA; 40622 MW; 2EB3947D CRC32;
```

```
Query Match 2.8%; Score 10; DB 1; Length 355;
Best Local Similarity 100.0%; Pred. No. 0.025;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 133 DRYLAIVHAT 142
    |||||
Db 139 DRYLAIVHAT 148
```

```
RESULT 15
ID IL8A_RAT STANDARD: PRT: 349 AA.
AC P70612;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE HIGH AFFINITY INTERLEUKIN-8 RECEPTOR A (IL-8R A) (CXCR-1).
GN IL8RA OR CXCR1.
OS RATIUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC EUROENTIA; SCIUROGNATHI; MORIDAE; MORINAE; RATIUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MISTAR; TISSUE=LUNG;
RC MEDLINE: 97115810.
RA DUNSTAN C.-A.N., SALAFRANCA M.N., ADHIKARI S., XIA Y., FENG L.,
RA HARRISON J.K.;
RT "Identification of two rat genes orthologous to the human
interleukin-8 receptors."
RT J. Biol. Chem. 271:32770-32776(1996).
RL -1- FUNCTION: RECEPTOR TO INTERLEUKIN-8, WHICH IS A POWERFUL
CC NEUTROPHILS CHEMOTACTIC FACTOR. BINDING OF IL-8 TO THE RECEPTOR
CC CAUSES ACTIVATION OF NEUTROPHILS. THIS RESPONSE IS MEDIATED VIA A
CC G-PROTEIN THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
```

```
CC MESSENGER SYSTEM (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U71089; G1589931; -
DR GCRDB: GCR_1404; -
DR PROSITE: PS00237; G_PROTEIN_RECEPTOR; 1.
DR PFAM: PF00001; 7tm_1; 1.
DR HSSP: P34996; 1DDD.
KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN;
KW CHEMOTAXIS.
FT DOMAIN 1 44
FT TRANSSEM 45 71
FT TRANSSEM 72 80
FT TRANSSEM 81 101
FT TRANSSEM 102 116
FT TRANSSEM 117 138
FT TRANSSEM 139 159
FT TRANSSEM 160 179
FT TRANSSEM 180 204
FT TRANSSEM 205 225
FT TRANSSEM 226 247
FT TRANSSEM 248 269
FT TRANSSEM 270 290
FT TRANSSEM 291 313
FT TRANSSEM 314 349
FT CARBOHYD 22 22
FT DISULFID 115 192
FT SEQUENCE 349 AA; 39944 MW; A8415DD5 CRC32;
```

```
Query Match 2.8%; Score 10; DB 1; Length 349;
Best Local Similarity 100.0%; Pred. No. 0.024;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 133 DRYLAIVHAT 142
    |||||
Db 139 DRYLAIVHAT 148
```

```
RESULT 16
ID IL8B_BOVIN STANDARD: PRT: 360 AA.
AC Q28003;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE HIGH AFFINITY INTERLEUKIN-8 RECEPTOR B (IL-8R B) (CXCR-2).
GN IL8RB OR CXCR2.
OS BOS TAURUS (BOVINE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC ARTIODACTYLA; RUMINANTIA; PECORA; BOVIDAE; BOVIDAE; BOS.
RN [1]
RP SEQUENCE FROM N.A.
RL LI Y., FENG J., TEMPLETON J.W.;
RL SUBMITTED (JAN-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- FUNCTION: RECEPTOR TO INTERLEUKIN-8, WHICH IS A POWERFUL
CC NEUTROPHILS CHEMOTACTIC FACTOR. BINDING OF IL-8 TO THE RECEPTOR
CC CAUSES ACTIVATION OF NEUTROPHILS. THIS RESPONSE IS MEDIATED VIA A
CC G-PROTEIN THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
CC MESSENGER SYSTEM. THIS RECEPTOR BINDS TO IL-8 WITH A HIGH AFFINITY
CC AND TO GRO/MGSA AND MAP-2 ALSO WITH A HIGH AFFINITY (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
```



```
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U19947; G1145878;
DR GCRDB: GCR_1224;
DR PROSITE: PS00237; G-PROTEIN_RECEPTOR; 1.
DR PFAM: PF00001; 7tm_1; 1.
KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN;
KW CHEMOTAXIS.
FT DOMAIN 1 48 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 49 75 1 (POTENTIAL).
FT DOMAIN 76 84 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 85 105 2 (POTENTIAL).
FT DOMAIN 106 120 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 121 142 3 (POTENTIAL).
FT DOMAIN 143 163 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 164 183 4 (POTENTIAL).
FT DOMAIN 184 208 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 209 231 5 (POTENTIAL).
FT DOMAIN 232 251 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 252 273 6 (POTENTIAL).
FT DOMAIN 274 294 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 295 315 7 (POTENTIAL).
FT DOMAIN 316 360 CYTOPLASMIC (POTENTIAL).
FT DISULFID 119 196 BY SIMILARITY.
FT CARBOHYD 10 10 POTENTIAL.
FT CARBOHYD 24 24 POTENTIAL.
SQ SEQUENCE 360 AA; 40625 MW; 281DD168 CRC32;

Query Match 2.8%; Score 10; DB 1; Length 360;
Best Local Similarity 100.0%; Pred. No. 0.025;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 133 DRYLAIVHAT 142
Db 143 DRYLAIVHAT 152

RESULT 17
IL8B_GORGO
ID IL8B_GORGO STANDARD; PRT; 353 AA.
AC Q28422;
DT 01-NOV-1997 (REL. 35, CREATED)
DI 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DI 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE HIGH AFFINITY INTERLEUKIN-8 RECEPTOR B (IL-8R B) (CXCR-2) (FRAGMENT).
GN IL8RB OR CXCR2.
OS GORILLA GORILLA GORILLA (LOWLAND GORILLA).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; GORILLA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 96175151.
RA ALVAREZ V., COTO E., SETIEN F., GONZALEZ S., GONZALEZ-ROCES S.,
RA LOPEZ-LARREA C.;
RT "Characterization of interleukin-8 receptors in non-human primates.";
RL IMMUNOGENETICS 43:261-267(1995).
CC -1- FUNCTION: RECEPTOR TO INTERLEUKIN-8, WHICH IS A POWERFUL
CC NEUTROPHILS CHEMOTACTIC FACTOR. BINDING OF IL-8 TO THE RECEPTOR
CC CAUSES ACTIVATION OF NEUTROPHILS. THIS RESPONSE IS MEDIATED VIA A
CC G-PROTEIN THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
CC MESSENGER SYSTEM. THIS RECEPTOR BINDS TO IL-8 WITH A HIGH AFFINITY
CC AND TO GRO/MGSA AND NAP-2 ALSO WITH A HIGH AFFINITY.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
```

```
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X91114; E198175;
DR PROSITE: PS00237; G-PROTEIN_RECEPTOR; 1.
DR PFAM: PF00001; 7tm_1; 1.
DR HSSP: P34996; 1DD;
KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN;
KW CHEMOTAXIS.
FT NON_TER 1 1
FT DOMAIN <1 45 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 46 72 1 (POTENTIAL).
FT DOMAIN 73 81 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 82 102 2 (POTENTIAL).
FT DOMAIN 103 117 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 118 139 3 (POTENTIAL).
FT DOMAIN 140 160 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 161 180 4 (POTENTIAL).
FT DOMAIN 181 205 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 206 228 5 (POTENTIAL).
FT DOMAIN 229 248 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 249 270 6 (POTENTIAL).
FT DOMAIN 271 291 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 292 312 7 (POTENTIAL).
FT DOMAIN 313 >353 CYTOPLASMIC (POTENTIAL).
FT DISULFID 116 193 BY SIMILARITY.
FT CARBOHYD 19 19 POTENTIAL.
FT NON_TER 353 353
SQ SEQUENCE 353 AA; 39919 MW; 4AF43313 CRC32;
```

```
Query Match 2.8%; Score 10; DB 1; Length 353;
Best Local Similarity 100.0%; Pred. No. 0.025;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 133 DRYLAIVHAT 142
Db 140 DRYLAIVHAT 149

RESULT 18
IL8B_HUMAN
ID IL8B_HUMAN STANDARD; PRT; 360 AA.
AC P25025;
DT 01-MAY-1992 (REL. 22, CREATED)
DI 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DI 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE HIGH AFFINITY INTERLEUKIN-8 RECEPTOR B (IL-8R B) (CXCR-2) (GRO/MGSA
DE RECEPTOR) (IL-8 RECEPTOR TYPE 2).
GN IL8RB OR CXCR2.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 91368200.
RA MURPHY P.M., TIFFANY H.L.;
RT "Cloning of complementary DNA encoding a functional human
RT interleukin-8 receptor.";
RL SCIENCE 253:1280-1283(1991).
CC [2]
CC SEQUENCE FROM N.A., AND CHARACTERIZATION.
CC MEDLINE; 93205012.
CC CERRETTI D.P., KOZLOSKY C.J., VANDEN BOS T., NELSON N., GEARING D.P.,
CC BECKMANN M.P.;
CC "Molecular characterization of receptors for human interleukin-8,
CC GRO/melanoma growth-stimulatory activity and neutrophil activating
CC peptide-2.";
RT
```


FT CARBOHYD 19 19 POTENTIAL.
 FT NON_TER 353 353
 SO SEQUENCE 353 AA: 39947 MW: 8592FE64 CRC32:

Query Match 2.8%: Score 10; DB 1; Length 353;
 Best Local Similarity 100.0%: Pred. No. 0.025;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 133 DRYLAIVHAT 142
 |||||
 Db 140 DRYLAIVHAT 149

RESULT 20
 IL8B_MOUSE STANDARD; PRT: 359 AA.
 AC P35343;
 DT 01-JUN-1994 (REL. 29, CREATED)
 DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
 DE HIGH AFFINITY INTERLEUKIN-8 RECEPTOR B (IL-8R B) (CXCR-2) (GRO/MSA
 DE RECEPTOR).
 GN IL8RB OR CXCR2 OR CXKR2 OR GPCR16.
 OS MUS MUSCULUS (MOUSE).
 OC EUKARYOTA: METAZOA: CHORDATA: VERTEBRATA: MAMMALIA: EUTHERIA:
 OC RODENTIA: SCIOGOGNATHI: MORIDAE: MORINAE: MUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 94117014.
 RA CERRETTI D.P., NELSON N., KOZLOSKY C.J., MORRISSEY P.J.,
 RA COPELAND N.G., GILBERT D.J., JENKINS N.A., DOSIK J.K., MOCK B.A.,
 RT "The murine homologue of the human interleukin-8 receptor type B maps
 near the Ily-1sh-Bog disease resistance locus.";
 RL GENOMICS 18:410-413(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/C;
 RX MEDLINE: 95050765.
 RA BOZIC C.R., GERARD N.P., VON UEXKULL-GULDENBAND C.,
 RA KOLAKOWSKI L.F. JR., CONKLYN M.J., BRESLOW R., SHOWELL H.J.,
 RA GERARD C.,
 RT "The murine interleukin 8 type B receptor homologue and its ligands.
 RT Expression and biological characterization.";
 RL J. BIOL. CHEM. 269:29355-29358(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-129/SV;
 RX MEDLINE: 95363183.
 RA LEE J., CACALANO G., CAMERATO T., TOY K., MOORE M.W., WOOD W.I.,
 RT "Chemokine binding and activities mediated by the mouse IL-8
 RT receptor.";
 RL J. IMMUNOL. 155:2158-2164(1995).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 94252584.
 RA HARADA A., KUNO K., NOMURA H., MUKAIDA N., MURAKAMI S., MATSUSHIMA K.,
 RT "Cloning of a cDNA encoding a mouse homolog of the interleukin-8
 RT receptor.";
 RL GENE 142:297-300(1994).
 RN [5]
 RP SEQUENCE OF 145-258 FROM N.A.
 RC TISSUE-TESTIS;
 RX MEDLINE: 94116980.
 RA WILKIE T.M., CHEN Y., GILBERT D.J., MOORE K.J., YU L., SIMON M.I.,
 RA COPELAND N.G., JENKINS N.A.,
 RT "Identification, chromosomal location, and genome organization of
 RT mammalian G-protein-coupled receptors.";
 RL GENOMICS 18:175-184(1993).
 CC -1- FUNCTION: RECEPTOR TO INTERLEUKIN-8, WHICH IS A POWERFUL
 CC NEUTROPHILS CHEMOTACTIC FACTOR. BINDING OF IL-8 TO THE RECEPTOR
 CC CAUSES ACTIVATION OF NEUTROPHILS. THIS RESPONSE IS MEDIATED VIA A
 CC G-PROTEIN THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND

CC MESSENGER SYSTEM. THIS RECEPTOR BINDS TO IL-8 WITH A HIGH AFFINITY
 CC AND TO GRO/MSA AND NAP-2 ALSO WITH A HIGH AFFINITY.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L23637; G435094; -
 DR EMBL: L13239; G293666; -
 DR EMBL: U31207; G950175; -
 DR EMBL: D17630; G493672; -
 DR EMBL: L20337; G438801; -
 DR PIR: A53677; A53677.
 DR GCRDB: GCR_0550; -
 DR GCRDB: GCR_0812; -
 DR GCRDB: GCR_0838; -
 DR GCRDB: GCR_1027; -
 DR GCRDB: GCR_1670; -
 DR MGD: MGI:105303; CXKR2.
 DR PROSITE: PS00237; G-PROTEIN RECEPTOR: 1.
 DR PFAM: PF00001; 7tm_1; 1.
 KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN;
 KW CHEMOTAXIS.
 FT DOMAIN 1 47
 FT TRANSEM 48 74
 FT DOMAIN 75 83
 FT TRANSEM 84 104
 FT DOMAIN 105 119
 FT TRANSEM 120 141
 FT DOMAIN 142 162
 FT TRANSEM 163 182
 FT DOMAIN 183 207
 FT TRANSEM 208 230
 FT DOMAIN 231 250
 FT TRANSEM 251 272
 FT DOMAIN 273 293
 FT TRANSEM 294 314
 FT DOMAIN 315 359
 FT CARBOHYD 223 223
 FT DISULFID 118 195
 SO SEQUENCE 359 AA: 40425 MW: D501949A CRC32:

Query Match 2.8%: Score 10; DB 1; Length 359;
 Best Local Similarity 100.0%: Pred. No. 0.025;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 133 DRYLAIVHAT 142
 |||||
 Db 142 DRYLAIVHAT 151

RESULT 21
 IL8B_PANTR STANDARD; PRT: 353 AA.
 AC 028807;
 DT 01-NOV-1997 (REL. 35, CREATED)
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE HIGH AFFINITY INTERLEUKIN-8 RECEPTOR B (IL-8R B) (CXCR-2) (FRAGMENT).
 GN IL8RB OR CXCR2.
 OS PAN TROGLODYTES (CHIMPANZEE).
 OC EUKARYOTA: METAZOA: CHORDATA: VERTEBRATA: MAMMALIA: EUTHERIA:
 OC PRIMATES: CATARRHINI: HOMINIDAE: PAN.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 96175151.

RA ALVAREZ V., COTO E., SETIEN F., GONZALEZ S., GONZALEZ-ROCES S.,
 RA LOPEZ-LARREA C.:
 "Characterization of Interleukin-8 receptors in non-human primates."
 RL IMMUNOGENETICS 43:261-267(1996).
 CC -1- FUNCTION: RECEPTOR TO INTERLEUKIN-8, WHICH IS A POWERFUL
 CC NEUTROPHILS CHEMOTACTIC FACTOR. BINDING OF IL-8 TO THE RECEPTOR
 CC CAUSES ACTIVATION OF NEUTROPHILS. THIS RESPONSE IS MEDIATED VIA A
 CC G-PROTEIN THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
 CC MESSENGER SYSTEM. THIS RECEPTOR BINDS TO IL-8 WITH A HIGH AFFINITY
 CC AND TO GRO/MGSA AND NAP-2 ALSO WITH A HIGH AFFINITY.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X91113; E198174; -
 DR PROSITE: PS00237; G_PROTEIN_RECEPTOR; 1.
 DR PFAM: PF00001; 7tm_1; 1.
 DR HSSP: P34996; 1DDD.
 KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN;
 KM CHEMOTAXIS.
 FT NON_TER 1 1
 FT DOMAIN <1 45 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 46 72 1 (POTENTIAL).
 FT DOMAIN 73 81 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 82 102 2 (POTENTIAL).
 FT DOMAIN 103 117 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 118 139 3 (POTENTIAL).
 FT DOMAIN 140 160 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 161 180 4 (POTENTIAL).
 FT DOMAIN 181 205 5 (POTENTIAL).
 FT TRANSMEM 206 228 6 (POTENTIAL).
 FT DOMAIN 229 248 7 (POTENTIAL).
 FT TRANSMEM 249 270 8 (POTENTIAL).
 FT DOMAIN 271 291 9 (POTENTIAL).
 FT TRANSMEM 292 312 10 (POTENTIAL).
 FT DOMAIN 313 353 11 (POTENTIAL).
 FT DISULFID 116 193 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 19 19 BY SIMILARITY.
 FT NON_TER 353 353 POTENTIAL.
 SO SEQUENCE 353 AA; 39998 MW; F4564B58 CRC32;

Query Match 2.8%; Score 10; DB 1; Length 353;
 Best Local Similarity 100.0%; Pred. No. 0.025;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 133 DRYLAIVHAT 142
 DB 140 DRYLAIVHAT 149

RESULT 22
 ID IL8B_RABIT STANDARD; PRT; 358 AA.
 AC P35344;
 DT 01-JUN-1994 (REL. 29, CREATED)
 DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
 DE 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE HIGH AFFINITY INTERLEUKIN-8 RECEPTOR B (IL-8R B) (CXCR-2) (GRO/MGSA
 DE RECEPTOR).
 GN IL8RB OR CXCR2.
 OS ORYCTOLAGUS Cuniculus (RABBIT).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
 OC LAGOMORPHA; LEPORIDAE; ORYCTOLAGUS.
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN-ALBINO; TISSUE-BLOOD;
 RX MEDLINE: 94230294.
 RA PRADO G.N., THOMAS K.M., SUZUKI H., LAROSA G.J., WILKINSON N.C.,
 RA FOLCO E., NAVARRO J.;
 RT "Molecular characterization of a novel rabbit interleukin-8 receptor
 RT isotype".
 RL J. Biol. Chem. 269:12391-12394(1994).
 CC -1- FUNCTION: RECEPTOR TO INTERLEUKIN-8, WHICH IS A POWERFUL
 CC NEUTROPHILS CHEMOTACTIC FACTOR. BINDING OF IL-8 TO THE RECEPTOR
 CC CAUSES ACTIVATION OF NEUTROPHILS. THIS RESPONSE IS MEDIATED VIA A
 CC G-PROTEIN THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
 CC MESSENGER SYSTEM. THE AFFINITY OF THIS RECEPTOR IS IL-8 >> NAP-2 >
 CC MGSA (GRO).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: EXPRESSED PREFERENTIALLY IN NEUTROPHILS.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: L24445; G437662; -
 DR PIR: A53752; A53752.
 DR GDB: GCR-0861; -
 DR PROSITE: PS00237; G_PROTEIN_RECEPTOR; 1.
 DR PFAM: PF00001; 7tm_1; 1.
 KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN;
 KM CHEMOTAXIS.
 FT DOMAIN 1 46 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 47 73 1 (POTENTIAL).
 FT DOMAIN 74 82 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 83 103 2 (POTENTIAL).
 FT DOMAIN 104 118 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 119 140 3 (POTENTIAL).
 FT DOMAIN 141 161 4 (POTENTIAL).
 FT TRANSMEM 162 181 5 (POTENTIAL).
 FT DOMAIN 182 206 6 (POTENTIAL).
 FT TRANSMEM 207 229 7 (POTENTIAL).
 FT DOMAIN 230 249 8 (POTENTIAL).
 FT TRANSMEM 250 271 9 (POTENTIAL).
 FT DOMAIN 272 292 10 (POTENTIAL).
 FT TRANSMEM 293 313 11 (POTENTIAL).
 FT CARBOHYD 314 358 12 (POTENTIAL).
 FT NON_TER 8 8 POTENTIAL.
 FT CARBOHYD 8 20 POTENTIAL.
 FT DISULFID 117 194 BY SIMILARITY.
 SO SEQUENCE 358 AA; 40632 MW; 90C46E08 CRC32;

Query Match 2.8%; Score 10; DB 1; Length 358;
 Best Local Similarity 100.0%; Pred. No. 0.025;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 133 DRYLAIVHAT 142
 DB 141 DRYLAIVHAT 150

RESULT 23
 ID IL8B_RAT STANDARD; PRT; 359 AA.
 AC P35407;
 DT 01-JUN-1994 (REL. 29, CREATED)
 DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
 DE HIGH AFFINITY INTERLEUKIN-8 RECEPTOR B (IL-8R B) (CXCR-2) (GRO/MGSA
 DE RECEPTOR).
 GN IL8RB OR CXCR2.
 OS RATIUS NORVEGICUS (RAT).

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
 OC RODENTIA; SCIOGOGNATHI; MORIDAE; MURINAE; RATTUS.
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=LUNG;
 RA GOBL A.E., WANG S., ZHOU Y., OEBERG K.;
 RL SUBMITTED (FEB-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN=WISTAR; TISSUE=LIVER;
 RA KONISHI K., SHIBATA F., WATANABE K., TSURUFUJI S., NAKAGAWA H.,
 RA FUJIOKA M.;
 RN SUBMITTED (JUL-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
 RP (3)
 RP SEQUENCE FROM N.A.
 RC STRAIN=WISTAR; TISSUE=SPLEEN, AND LUNG;
 RA MEDLINE: 97115810.
 RX DUNSTAN C., A.N., SALAFRANCA M.N., ADHIKARI S., XIA Y., FENG L.,
 RA HARRISON J.K.;
 RT Identification of two rat genes orthologous to the human
 RT interleukin-8 receptors.
 RL J. BIOL. CHEM. 271:32770-32776(1996).
 CC -1- FUNCTION: RECEPTOR TO INTERLEUKIN-8, WHICH IS A POWERFUL
 CC NEUTROPHILS CHEMOTACTIC FACTOR. BINDING OF IL-8 TO THE RECEPTOR
 CC CAUSES ACTIVATION OF NEUTROPHILS. THIS RESPONSE IS MEDIATED VIA A
 CC G-PROTEIN THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
 CC MESSENGER SYSTEM. THIS RECEPTOR BINDS TO IL-8 WITH A HIGH AFFINITY
 CC AND TO GRO/MGSA AND NAP-2 ALSO WITH A HIGH AFFINITY.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X77797; G498703; -
 DR EMBL: D63584; G944819; -
 DR EMBL: U70988; G1617613; -
 DR PIR: S42096; S42096.
 DR GCRDB: GCR_0913; -
 DR GCRDB: GCR_1405; -
 DR GCRDB: GCR_1524; -
 DR PROSITE: PS00237; G-PROTEIN_RECEPTOR; 1.
 DR PFAM: PF00001; 7tm_1; 1.
 KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN;
 KW CHEMOTAXIS.
 FT DOMAIN 1 47
 FT TRANSMEM 48 74
 FT DOMAIN 75 83
 FT TRANSMEM 84 104
 FT DOMAIN 105 119
 FT TRANSMEM 120 141
 FT DOMAIN 142 162
 FT TRANSMEM 163 182
 FT DOMAIN 183 207
 FT TRANSMEM 208 230
 FT DOMAIN 231 250
 FT TRANSMEM 251 272
 FT DOMAIN 273 296
 FT TRANSMEM 297 314
 FT DOMAIN 315 358
 FT CARBOHYD 8 8
 FT CARBOHYD 23 23
 FT CARBOHYD 201 201
 FT CARBOHYD 202 202
 FT DISULFID 118 195
 FT SEQUENCE 359 AA; 40532 MW; 5829D194 CRC32;

Query Match 2.88; Score 10; DB 1; Length 359;
 Best Local Similarity 100.0%; Pred. No. 0.025;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 133 DRYLAIVHAT 142
 |||||||||
 Db 142 DRYLAIVHAT 151
 RESULT 24
 OPRX_CAVPO STANDARD; PRT; 370 AA.
 ID OPRX_CAVPO
 AC P47748;
 DT 01-FEB-1996 (REL. 33, CREATED)
 DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE NOCICEPTIN RECEPTOR (ORPHANIN FQ RECEPTOR) (KAPPA-TYPE 3 OPIOID
 DE RECEPTOR (KOR-3) (XOR).
 GN OPR1 OR OOR.
 OS CAVIA PORCELLUS (GUINEA PIG).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
 OC RODENTIA; HYSTRICOGNATHI; CAVIIDAE; CAVIA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HARTLEY; TISSUE=BRAIN.
 RA XIE G.;
 RL SUBMITTED (DEC-1993) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC -1- FUNCTION: RECEPTOR FOR THE NEUROPEPTIDE NOCICEPTIN/ORPHANIN FQ.
 CC HAS A POTENTIAL ROLE IN MODULATING A NUMBER OF BRAIN FUNCTIONS,
 CC INCLUDING INSTINCTIVE BEHAVIORS AND EMOTIONS. THE ACTIVITY OF
 CC THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH INHIBITS ADENYLYL
 CC CYCLASE.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U04369; G606789; -
 DR GCRDB: GCR_1169; -
 DR PROSITE: PS00237; G-PROTEIN_RECEPTOR; 1.
 DR PFAM: PF00001; 7tm_1; 1.
 DR HSP: P34996; 1DD.
 KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN;
 KW PHOSPHORYLATION; LIPOPROTEIN; PALMITATE.
 FT DOMAIN 1 50
 FT TRANSMEM 51 77
 FT DOMAIN 78 87
 FT TRANSMEM 88 109
 FT DOMAIN 110 124
 FT TRANSMEM 125 146
 FT DOMAIN 147 165
 FT TRANSMEM 166 188
 FT DOMAIN 189 211
 FT TRANSMEM 212 236
 FT DOMAIN 237 264
 FT TRANSMEM 265 288
 FT DOMAIN 289 300
 FT TRANSMEM 301 322
 FT DOMAIN 323 370
 FT CARBOHYD 123 200
 FT LIPID 334 334
 FT CARBOHYD 21 21
 FT CARBOHYD 28 28
 FT CARBOHYD 39 39
 FT SEQUENCE 370 AA; 40789 MW; 4ADFAEEF CRC32;

RE J. NEUROINFORM. 59:91-101(1995).
CC -1- FUNCTION: RECEPTOR FOR THE NEUROPEPTIDE NOCICEPTIN/ORPHANIN FQ.
CC HAS A POTENTIAL ROLE IN MODULATING A NUMBER OF BRAIN FUNCTIONS,
CC INCLUDING INSTINCTIVE BEHAVIORS AND EMOTIONS. THE ACTIVITY OF
CC THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH INHIBITS ADENYLYL
CC CYCLASE.
CC
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U04952; G440880; -
CC EMBL: D31657; G808874; -
CC EMBL: D31655; G808874; JOINED.
CC EMBL: X91813; G1008982; -
CC EMBL: U32932; G1464791; -
CC EMBL: U32928; G1464791; JOINED.
CC EMBL: U32930; G1464791; JOINED.
CC EMBL: U09421; G551485; -
CC EMBL: U14165; G540093; -
CC GCRDB: GCR_0891; -
CC GCRDB: GCR_1715; -
CC GCRDB: GCR_1715; -
CC GCRDB: GCR_1728; -
CC GCRDB: GCR_1731; -
CC MGD: MGI:97440; OPRL.
CC PROSITE: PS00237; G-PROTEIN-RECEPTOR: 1.
CC PFM: PFM0001; 7em_1; 1.
CC HSSP: P34996; 1DD0.
CC
CC G-PROTEIN COUPLED RECEPTOR: TRANSMEMBRANE; GLYCOPROTEIN;
CC PHOSPHORYLATION: LIPOPROTEIN; PALMITATE.
CC
CC DOMAIN 1 47
CC TRANSMEM 48 74
CC TRANSMEM 75 84
CC TRANSMEM 85 106
CC TRANSMEM 107 121
CC TRANSMEM 122 143
CC TRANSMEM 144 162
CC TRANSMEM 163 185
CC TRANSMEM 186 208
CC TRANSMEM 209 233
CC TRANSMEM 234 261
CC TRANSMEM 262 285
CC TRANSMEM 286 297
CC TRANSMEM 298 319
CC TRANSMEM 320 365
CC TRANSMEM 365 197
CC TRANSMEM 120 331
CC TRANSMEM 331 331
CC TRANSMEM 21 21
CC TRANSMEM 26 26
CC TRANSMEM 36 36
CC TRANSMEM 348 349
CC TRANSMEM 367 AA; 40491 MM; 3f472156 CRC32;
CC
CC SEQUENCE

Query Match 2.8%; Score 10; DB 1; Length 367;
Best Local Similarity 100.0%; Pred. No. 0.025;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 296 CLNPITVAF 305
DB 310 CLNPITVAF 319

RESULT 27
OPRX_PIG STANDARD: PRT: 370 AA.

AC P79292;
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DE NOCICEPTIN RECEPTOR (ORPHANIN FQ RECEPTOR) (KAPPA-TYPE 3 OPIOID
DE RECEPTOR) (KOR-3) (ORGC) (K3 OPIATE RECEPTOR).
GN OPRL.
OS SUS SCROFA (PIG).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC ARTIODACTYLA; SUIFORMES; SUINA; SUIDAE; SUS.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN CORTEX;
RA OSINSKI M.A., PAMPUSCH M.S., BROWN D.R., MURTAGH M.P.;
RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DDAY DATA BANKS.
CC -1- FUNCTION: RECEPTOR FOR THE NEUROPEPTIDE NOCICEPTIN/ORPHANIN FQ.
CC HAS A POTENTIAL ROLE IN MODULATING A NUMBER OF BRAIN FUNCTIONS,
CC INCLUDING INSTINCTIVE BEHAVIORS AND EMOTIONS. THE ACTIVITY OF
CC THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH INHIBITS ADENYLYL
CC CYCLASE (BY SIMILARITY).
CC
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U72758; G1763257; -
CC GCRDB: GCR_1519; -
CC PROSITE: PS00237; G-PROTEIN-RECEPTOR: 1.
CC PFM: PFM0001; 7em_1; 1.
CC HSSP: P34996; 1DD0.
CC
CC G-PROTEIN COUPLED RECEPTOR: TRANSMEMBRANE; GLYCOPROTEIN;
CC PHOSPHORYLATION: LIPOPROTEIN; PALMITATE.
CC
CC DOMAIN 1 50
CC TRANSMEM 51 77
CC TRANSMEM 78 87
CC TRANSMEM 88 109
CC TRANSMEM 110 124
CC TRANSMEM 125 146
CC TRANSMEM 147 165
CC TRANSMEM 166 188
CC TRANSMEM 189 211
CC TRANSMEM 212 236
CC TRANSMEM 237 264
CC TRANSMEM 265 288
CC TRANSMEM 289 300
CC TRANSMEM 301 322
CC TRANSMEM 323 370
CC TRANSMEM 370 200
CC TRANSMEM 123 334
CC TRANSMEM 334 334
CC TRANSMEM 21 21
CC TRANSMEM 28 28
CC TRANSMEM 39 39
CC TRANSMEM 370 AA; 40610 MM; E1050DAC CRC32;
CC
CC SEQUENCE

Query Match 2.8%; Score 10; DB 1; Length 370;
Best Local Similarity 100.0%; Pred. No. 0.026;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 296 CLNPITVAF 305
DB 313 CLNPITVAF 322

RESULT 28
OPRX_RAT STANDARD: PRT: 367 AA.

AC P35370;
 DT 01-JUN-1994 (REL. 29, CREATED)
 DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
 DE NOCICEPTIN RECEPTOR (ORPHANIN FQ RECEPTOR) (KAPPA-TYPE 3 OPIOID
 RECEPTOR) (KOR-3) (FOR-C) (XOR1).
 GN OPRL1 OR OOR.
 OS RATTUS NORVEGICUS (RAT).
 OC EURKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
 OC RODENTIA; SCIUROGNATHI; MORIDAE; MORINAE; RATTUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MISTAR; TISSUE=BRAIN;
 RX MEDLINE: 94215703.
 RA FUKUDA K., KATO S., MORI K., NISHI M., TAKESHIMA H., INABE N.,
 RA MIYATA T., HOUTANI T., SUGIMOTO T.,
 RT "CDNA cloning and regional distribution of a novel member of the
 RT opioid receptor family.";
 RL FEBS LETT. 343:42-46(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=HIPPOCAMPUS;
 RA MENG F., XIE G., ALFRED M., THOMPSON R., HOVERSTEN M., WATSON S.,
 RA AKIL H.;
 RL SUBMITTED (JAN-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=BRAIN;
 RX MEDLINE: 94307401.
 RA BUNZOW J.R., SAEZ C., MORTIUD M., BOUVIER C., WILLIAMS J.T., LOW M.,
 RA GRANDY D.K.;
 RT "Molecular cloning and tissue distribution of a putative member of
 RT the rat opioid receptor gene family that is not a mu, delta or kappa
 RT opioid receptor type.";
 RL FEBS LETT. 347:284-288(1994).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC MEDLINE: 94307400.
 RA CHEN Y., FAN Y., LU J., MESTEK A., TIAN M., KOZAK C.A., YU L.;
 RT "Molecular cloning, tissue distribution and chromosomal localization
 RT of a novel member of the opioid receptor gene family.";
 RL FEBS LETT. 347:279-283(1994).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BRAIN;
 RX MEDLINE: 95096849.
 RA LACHOWICZ J.E., SHEN Y., MONSMA F.J. JR., SIBLEY D.R.;
 RT "Molecular cloning of a novel G protein-coupled receptor related to
 RT the opiate receptor family.";
 RL J. NEUROCHEM. 64:34-40(1995).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BRAIN;
 RX MEDLINE: 94298959.
 RA WANG J.B., JOHNSON P.S., IMAI Y., PERSICO A.M., OZENBERGER B.A.,
 RA EPPLER C.M., UHL G.R.;
 RT "CDNA cloning of an orphan opiate receptor gene family member and its
 RT splice variant.";
 RL FEBS LETT. 348:75-79(1994).
 RN [7]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=BRAIN;
 RX MEDLINE: 95182817.
 RA WICK M.J., MINNERATH S.R., LIN X., ELDE R.P., LAW P.Y., LOH H.H.;
 RT "Isolation of a novel cDNA encoding a putative membrane receptor with
 RT high homology to the cloned mu, delta, and kappa opioid receptors.";
 RL BRAIN RES. MOL. BRAIN RES. 27:37-44(1994).
 CC -1- FUNCTION: RECEPTOR FOR THE NEUROPEPTIDE NOCICEPTIN/ORPHANIN FQ.
 CC HAS A POTENTIAL ROLE IN MODULATING A NUMBER OF BRAIN FUNCTIONS,
 CC INCLUDING INSTINCTIVE BEHAVIORS AND EMOTIONS. THE ACTIVITY OF
 CC THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH INHIBITS ADENYLYL
 CC CYCLASE.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.

CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN SEVERAL BRAIN AREAS.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D16438; G533355; -
 DR EMBL: U05239; G451844; -
 DR EMBL: U01913; G487965; -
 DR EMBL: L28144; G496220; -
 DR EMBL: U07871; G606803; -
 DR EMBL: L33916; G557200; -
 DR EMBL: L29419; G510719; -
 DR PIR: S46238; S46238.
 DR PIR: S43655; S43655.
 DR GCRDB: GCR_0834; -
 DR GCRDB: GCR_0898; -
 DR GCRDB: GCR_0912; -
 DR GCRDB: GCR_1030; -
 DR GCRDB: GCR_1455; -
 DR GCRDB: GCR_1487; -
 DR GCRDB: GCR_1517; -
 DR PROSITE: PS00237; G-PROTEIN_RECEPTOR; 1.
 DR PRAM: PF00001; 7tm_1; 1.
 DR HSP: P34996; 1DD0.
 KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN;
 KM PHOSPHORYLATION; LIPOPROTEIN; PALMITATE.
 FT DOMAIN 1 47
 FT TRANSMEM 48 74
 FT DOMAIN 75 84
 FT TRANSMEM 85 106
 FT DOMAIN 107 121
 FT TRANSMEM 122 143
 FT DOMAIN 144 162
 FT TRANSMEM 163 185
 FT DOMAIN 186 208
 FT TRANSMEM 209 233
 FT DOMAIN 234 261
 FT TRANSMEM 262 285
 FT DOMAIN 286 297
 FT TRANSMEM 298 319
 FT DOMAIN 320 366
 FT DISULFID 120 197
 FT LIPID 331 331
 FT CARBOHYD 21 21
 FT CARBOHYD 26 26
 FT CARBOHYD 36 36
 FT CONFLICT 105 105
 FT CONFLICT 226 226
 FT CONFLICT 246 246
 FT CONFLICT 348 348
 SQ SEQUENCE 367 AA; 40523 MW; 7FD40CCC CRC32;

Query Match 2.88; Score 10; DB 1; Length 367;
 Best Local Similarity 100.0%; Pred. No. 0.025;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 296 CLNPITVAF 305
 |||||
 Db 310 CLNPITVAF 319

Search completed: August 8, 1999, 06:07:34
 Job time: 501 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 8, 1999, 02:04:38 ; Search time 17.77 Seconds
(without alignments)
1219.097 Million cell updates/sec

Title: US-09-104-063-4
Perfect score: 352
Sequence: 1 MEGSITSDNYTEMGSGD.....KRGHSSVSTESSSFHSS 352

Scoring table: OLIGO

Searched: 201082 seqs, 61543640 residues

Database :

SPTREMBL_10.*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_undefined:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	347	98.6	350	050835	050835 homo sapien
2	176	50.0	352	077488	077488 cercopithec
3	60	17.0	332	11	062973 ratius norv
4	39	11.1	358	13	09ygc3 xenopus lae
5	19	5.4	353	13	093247 cyprinus ca
6	16	4.5	357	13	042445 oncorhynch
7	10	2.8	356	6	097571 canis fam1

ALIGNMENTS

Result	ID	Score	Query Match	Length	DB ID	Description
1	050835	347	98.6	350	050835	050835 homo sapien
2	077488	176	50.0	352	077488	077488 cercopithec
3	062973	60	17.0	332	11	062973 ratius norv
4	09ygc3	39	11.1	358	13	09ygc3 xenopus lae
5	093247	19	5.4	353	13	093247 cyprinus ca
6	042445	16	4.5	357	13	042445 oncorhynch
7	097571	10	2.8	356	6	097571 canis fam1

RT "Genomic organization and promoter characterization of human CXCR4
RT gene":
RL FEBS Lett. 426:271-278(1998).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL: AJ224869; CAA12166.1; -
DR PFM: PFO0001; 7cm.1; 1
DR PROSITE: PS00237; G-PROTEIN_RECEPTOR; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
SQ SEQUENCE 360 AA; 40607 MW; 110EB9E6 CRC32;

Query Match 98.6%; Score 347; DB 4; Length 360;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 347; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 6 IYTSNDYTEMGSGDYSKPECFRENNANFNKIFLPTIYSIIFLTGIYNGILVMGY 65
DB 14 IYTSNDYTEMGSGDYSKPECFRENNANFNKIFLPTIYSIIFLTGIYNGILVMGY 73
QY 66 QKKLSMTDKYRLHLSVADLLFVITLPEWADAVANWYGNFLCAVHYIVNLYSSVL 125
DB 74 QKKLSMTDKYRLHLSVADLLFVITLPEWADAVANWYGNFLCAVHYIVNLYSSVL 133
QY 126 ILAFISLDRIYLAIVHATNSQRPRLKLAEKVYVGVWIPALLITIPDFIFANVSEADRYI 185
DB 134 ILAFISLDRIYLAIVHATNSQRPRLKLAEKVYVGVWIPALLITIPDFIFANVSEADRYI 193
QY 186 CDRFPYNDLVVYVFOHIMVGLIPGIYILSCYCIISKLSHGKQKRLKTTVILI 245
DB 194 CDRFPYNDLVVYVFOHIMVGLIPGIYILSCYCIISKLSHGKQKRLKTTVILI 253
QY 246 LAFPMCLPYIYIGISIDSTILLEITIKGCEFNHYHAKMISTEALAFHCCINPIIYATL 305
DB 254 LAFPMCLPYIYIGISIDSTILLEITIKGCEFNHYHAKMISTEALAFHCCINPIIYATL 313
QY 306 GAKFITSAGHATSVSRGSLKILSKRGHSSVSTESSSFHSS 352
DB 314 GAKFITSAGHATSVSRGSLKILSKRGHSSVSTESSSFHSS 360

RESULT 2
ID 077488 PRELIMINARY: PRT: 352 AA.
AC 077488:
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)
DE CXCR4 RECEPTOR.
OS Cercopithecus aethiops (Green monkey) (Givet).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Cercopithecidae; Cercopithecinae;
OC Chlorocebus.
RN [1]
RP SEQUENCE FROM N.A.
RA MURAYAMA Y., MATSUNAGA S., INOUE-MURAYAMA M.:
RT "cDNA sequence of African green monkey CXCR-4 chemokine receptor
RT gene." Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RL -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC EMBL: AB015943; BAA31327.1; -
DR PFM: PFO0001; 7cm.1; 1.
DR PROSITE: PS00237; G-PROTEIN_RECEPTOR; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
SQ SEQUENCE 352 AA; 39773 MW; BD6D5150 CRC32;

Query Match 50.0%; Score 176; DB 6; Length 352;
Best Local Similarity 100.0%; Pred. No. 7; 6e-168;
Matches 176; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 177 VSEADRYICDRFPYNDLVVYVFOHIMVGLIPGIYILSCYCIISKLSHGKQKRL 236

```

|||||
Db 177 VSEADRTICDFEYNDLVWVVFQFQIMVGLIPGIVILSCYCIISLXSHKHQRK 236
OY 237 AKTIVILLAFACWLPYIGISIDSFILLEIKOGCFEFTVHKWISITFALAFHCC 296
Db 237 AKTIVILLAFACWLPYIGISIDSFILLEIKOGCFEFTVHKWISITFALAFHCC 296
OY 257 LNPITYAFGAFKTSQAQHALTSVSRGSSLYKLKSGKRGHSSVSTESSESSPHSS 352
Db 257 LNPITYAFGAFKTSQAQHALTSVSRGSSLYKLKSGKRGHSSVSTESSESSPHSS 352

```

```

RESULT 3
ID 062973 PRELIMINARY: PRT: 332 AA.
AC 062973:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
DE CHEMOKINE RECEPTOR LCRI.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY;
RA WONG M., XIN W.W., DOMAN R.S.;
RL MOL. PSYCH. 0:0-0(0).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DE EMBL: U54791; AAB01981.1; -.
DR PFAM: PF00001; 7tm_1; 2.
DR PROSITE: PS00237; G_PROTEIN_RECEPTOR; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
SO SEQUENCE 332 AA; 37442 MW; C0EAB848 CRC32;

```

```

Query Match 17.0%; Score 60; DB 11; Length 332;
Best Local Similarity 100.0%; Pred. No. 7.1e-52;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 197 VVFOFQIMVGLIPGIVILSCYCIISLXSHKHQRKALKTIVILLAFACWLPY 256
Db 177 VVFOFQIMVGLIPGIVILSCYCIISLXSHKHQRKALKTIVILLAFACWLPY 236

```

RESULT 4

ID 09YGC3 PRELIMINARY: PRT: 358 AA.

AC 09YGC3:

DT 01-MAY-1999 (TREMBLrel. 10, Created)

DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)

DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)

DE CHEMOKINE RECEPTOR 4.

GN CXCR4.

OS Xenopus laevis (African clawed frog).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Amphibia; Batrachia; Anura;

OC Mesobatrachia; Pipidae; Pipidae; Xenopodidae; Xenopus.

RN

RP SEQUENCE FROM N.A.

RA MOEPPS B., KNOEFFLE K., BROWN M., KNOEHEL W., GIERSCHEK P.;

RT "Expression of the CXCR chemokine receptor 4 during early Xenopus

RT laevis embryogenesis: a possible role of chemokine receptors as

RT regulators of development and differentiation."

RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

DE EMBL: Y17895; CAAT6924.1; -.

DE EMBL: Y17894; CAAT6923.1; -.

DR PROSITE: PS00237; G_PROTEIN_RECEPTOR; 1.

DR Receptor; G-protein coupled receptor; Transmembrane; Glycoprotein.

KW SEQUENCE 358 AA; 40078 MW; E991FD1D CRC32;

```

Query Match 11.1%; Score 39; DB 13; Length 358;
Best Local Similarity 100.0%; Pred. No. 7.6e-31;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 107 FLCKAVHYIYTNLYSVLLAFISLDRLAIVHATNSQ 145
Db 111 FLCKAVHYIYTNLYSVLLAFISLDRLAIVHATNSQ 149

```

RESULT 5

ID 093247 PRELIMINARY: PRT: 353 AA.

AC 093247:

DT 01-NOV-1998 (TREMBLrel. 08, Created)

DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)

DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)

DE CXCR4.

OS Cyprinus carpio (Common carp).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;

OC Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes;

OC Cyprinidae; Cyprinidae; Cyprininae; Cyprinus.

RN

RP SEQUENCE FROM N.A.

RA FUJIKI K., NAKAO M., SHIN D., YANO T.;

RT "CDNA cloning of a carp homologue of mammalian CXCR4."

RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

DE EMBL: AB012310; BAA32797.1; -.

DR PFAM: PF00001; 7tm_1; 1.

DR PROSITE: PS00237; G_PROTEIN_RECEPTOR; 1.

KW G-protein coupled receptor; Transmembrane; Glycoprotein.

SO SEQUENCE 353 AA; 39633 MW; 23DD5347 CRC32;

```

Query Match 5.4%; Score 19; DB 13; Length 353;
Best Local Similarity 100.0%; Pred. No. 7.5e-11;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 119 NLYSVLLAFISLDRLA 137
Db 117 NLYSVLLAFISLDRLA 135

```

RESULT 6

ID 042445 PRELIMINARY: PRT: 357 AA.

AC 042445:

DT 01-JAN-1998 (TREMBLrel. 05, Created)

DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)

DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)

DE CXCR CHEMOKINE RECEPTOR.

OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;

OC Neopterygii; Teleostei; Euteleostei; Protacanthopterygii;

OC Salmoniformes; Salmonidae; Oncorhynchus.

RN

RP SEQUENCE FROM N.A.

RA DANIELS G.D., CHARLEMAGNE J., SECOMBES C.J.;

RT "Cloning and sequencing of a rainbow trout, *Oncorhynchus mykiss*,

RT chemokine receptor homolog."

RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

DE EMBL: AJ001039; CA04493.1; -.

DR PFAM: PF00001; 7tm_1; 1.

DR PROSITE: PS00237; G_PROTEIN_RECEPTOR; 1.

DR G-protein coupled receptor; Transmembrane; Glycoprotein.

KW SEQUENCE 357 AA; 39817 MW; 6D892E7E CRC32;

```

Query Match 4.5%; Score 16; DB 13; Length 357;

```

Best Local Similarity 100.0%; Pred. No. 7.5e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 73 TDXYRLHLSVADLFFV 88
|||||
DB 75 TDXYRLHLSVADLFFV 90

RESULT 7

097571
ID 097571 PRELIMINARY; PRT; 356 AA.
AC 097571;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, last sequence update)
DT 01-MAY-1999 (TREMBlrel. 10, last annotation update)
DE INTERLEUKIN-8 RECEPTOR.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Carnivora; Fissipedia; Canidae; Canis.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BEED BEAGLE;
RA CHANG Y.F., NOVOSIL V., CHANG C.F.;
RT "The isolation and sequence of canine interleukin-8 receptor
homolog.";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF047047; AAC98968.1; -.
KW Receptor
SQ SEQUENCE 356 AA; 40505 MW; C2B7961F CRC32;

Query Match 2.8%; Score 10; DB 6; Length 356;
Best Local Similarity 100.0%; Pred. No. 0.075;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 133 DRYLAIVHAT 142
|||||
DB 141 DRYLAIVHAT 150

Search completed: August 8, 1999, 02:45:06
Job time: 2428 sec

